Investigating gene regulatory networks underlying zebrafish pigment cell development

Petratou, Kleio

Award date:
2017

Awarding institution:
University of Bath

Investigating gene regulatory networks underlying zebrafish pigment cell development

Kleio Petratou

A thesis submitted for the degree of Doctor of Philosophy

University of Bath

Department of Biology and Biochemistry

September 2016

“Models in systems biology are not meant to be descriptions, pathetic descriptions, of nature; they are designed to be accurate descriptions of our pathetic thinking about nature”

J. Black 1989/
J. Gunawardena 2014

Table of contents

List of Figures

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Appendices

List of Tables

Acknowledgements

Firstly, I would like to express my sincere gratitude to my supervisor, Prof. Robert Kelsh, for his constant support, patience and motivation. His enthusiasm truly inspired me throughout my PhD.

I am very grateful to my second supervisor, Dr Hartmut Schwetlick, for his guidance, understanding and for maintaining his sense of humour while I familiarised myself with the mathematics part of this project.

Besides my mentors, I am thankful to our collaborators, Dr Andrea Rocco and Dr Gerardo Aquino, for their valuable mathematical input and the stimulating discussions, and to Dr James Lister for sharing the *tfec* morpholino and CRISPR sequences, as well as exciting data. Special thanks to Dr Tatiana Subkhankulova for providing me with experimental input, resources and data.

I thank everyone in the zebramouse lab, for their advice, willingness to troubleshoot and the fun coffee breaks. A big thanks goes to Marc Shedden for his expert fish care.

I am especially thankful to my colleagues and close friends, Karen and Deeya, as well as Ravi for being so incredibly supportive and full of understanding throughout my PhD and for proofreading my thesis. I would also like to thank John for all his patience and support during these past four years.

Last but not least, a huge thank you goes to my family, particularly to my parents, Kyriakos and Effie, for guiding and supporting me at every step. Without them, no part of this PhD would have been possible.

Abstract

The multipotent cells comprising the vertebrate neural crest (NC) generate an astonishing array of derivatives, including neuronal, skeletal and adrenal components and pigment cells. Zebrafish possess three chromatophores lineages, melanophores, iridophores and xanthophores, which lend themselves to investigating the complex gene regulatory networks (GRNs) underlying fate segregation of NC progenitors. Although the core GRN governing melanophore specification has been previously established, those guiding iridophore and xanthophore development remain elusive.

This study explores the iridophore specification GRN via a previously employed systems biology approach. Loss and gain of function experiments were used to derive a preliminary GRN (model A), mathematically modelled using a system of differential equations and computationally simulated to predict gene expression dynamics. Predictions were experimentally evaluated and testable hypotheses were derived to render simulations of the resulting model B consistent with experimental observations. Iterations of this process led to a more sophisticated model D, which accurately predicted observed expression dynamics.

Firstly, mutant phenotypes identified the transcription factors Sox10, Tfec and Mitfa and the receptor tyrosine kinase, Ltk, as key players for model A. Regulatory interactions were derived by analysing the spatiotemporal gene expression patterns on wild-type and mutant embryos by *in situ* hybridisation. Data indicated a *sox10*-dependent *tfec*/*ltk* positive feedback loop driving iridophore specification. Cycles of mathematical modelling and experimentation revealed important regulatory features, such as *sox10* maintenance throughout iridophore development and its cooperation with *tfec* to activate the differentiation gene, *pnp4a*. Candidate repressors of *mitfa,* a melanocyte-specific target of *sox10,* were investigated. Surprisingly, data presented challenge the reported role of Foxd3, an established *mitfa* repressor, in iridophores.

This study established the core GRN guiding specification of the NC-derived iridophore lineage using systems biology. Presented results reveal previously unacknowledged molecular mechanisms underlying fate choice and demonstrate the value of integrating experimental and mathematical approaches when investigating GRNs.

Abbreviations

Chemicals

ADA	Adenosine deaminase
APS	Ammonium persulfate
BSA	Bovine serum albumin
DEPC	Diethylpyrocarbonate
DHI(CA)	5,6-dihydroxyindole (-2-car-boxylic acid)
DMSO	Dimethyl sulfoxide
DOPA	l-3,4- dihydroxyphenylalanine
ENU	N-ethyl-n-nitrosourea
EtOH	Ethanol
FACS	Fluorescence activated cell sorting
GMP/AMP	Guanosine and adenosine monophosphate
H_2O_2	Hydrogen peroxide
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KOH	Potassium hydroxide
LiCl	Lithium chloride
MeOH	Methanol
$MgCl_2$	Magnesium chloride
$MgSO_4$	Magnesium sulfate
NaCl	Sodium Chloride
NaOH	Sodium hydroxide
NBT/BCIP	Nitro-blue tetrazolium and 5-bromo-4-chloro-3'-indolyphosphate
PNP	Purine nucleoside phosphorylase
PTU	1-phenyl-2-thiourea
TEMED	Tetramethylethylenediamine
XGAL	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

Gene/protein names

BMP	Bone Morphogenetic Protein
Cas9	CRISPR associated protein 9
c-cbl	Casitas B-lineage Lymphoma
dct	Dopachrome tautomerase
Dlx	Distal-less gene of Drosophila
DNase	Deoxyribonuclease
FGF	Fibroblast Growth Factor
FGF(R)	Fibroblast growth factor receptor
Flt3	Fms-like tyrosine kinase 3
fox(d3/i)	Forkhead box (d3/i)
Gbx2	Gastrulation Brain Homeobox 2
gch	GTP cyclohydrolase
hdac	Histone deacetylase
HEK	Human EPH (ephrin)-like kinase
HGF(R)	Hepatocyte growth factor (receptor)
HTK(L)	Hepatoma transmembrane kinase ligand
Islet1	Insulin gene enhancer protein transcription factor
ltk	Leucocyte tyrosine kinase
mitf(a)	Microphthalmia-associated transcription factor (a)
Msx	Muscle segment homeobox gene
npm	Nucleophosmin
otx	Orthodenticle homeobox 1
pax3/7	Paired box 3/7
pnp4a	Purine nucleoside phosphorylase 4a
prdm1a	Positive regulatory domain I-binding factor 1
RNase	Ribonuclease
RT	Reverse transcriptase

RTK	Receptor tyrosine kinase
Shh	Sonic hedhehog
sox	SRY (Sex Determining Region Y)-related HMG box
TF	Transcription factor
Tfap2	Transcription Factor activator protein-2
tfec	Transcription factor EC
Trk	Tropomyocin related kinase
tyr	Tyrosinase
tyrp1	Tyrosinase-related protein 1
xdh	Xanthine dehydrogenase
xod	Xanthine oxidase
zic	Zinc finger in the cerebellum

Miscellaneous

a.a.	Amino acid(s)
bHLH-LZ	Basic helix–loop–helix leucine zipper
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
BS	Blocking solution
cDNA	Complementary DNA
ChIP-seq	Chromatin immunoprecipitation and sequencing
cNC(C)	Cranial neural crest (cell)
CNS/PNS	Central/peripheral nervous system
CRISPR	Clustered regularly interspaced short palindromic repeats
DIC	Differential interference contrast
DIG	Digoxigenin
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide Triphosphate
dpf	Days post fertilisation
E5I5	Exon 5-intron 5
EDTA	Ethylenediaminetetra-acetic acid
EpiSCs	Epiblast stem cells
ESCs	Embryonic stem cells
F (0/1/2)	Filial (generation 0/1/2)
FFL	Feed-forward loops
GRN	Gene regulatory networks
gRNA	Guide RNA
HDR	Homology-directed repair
HM	Hybridisation mix
HMG	High mobility group box
hpf	Hours post fertilisation
HRMA	High Resolution Melt Analysis
HSC	Haematopoietic stem cell
I/M	Iridoblast/melanoblast
ICM	Intermediate cell mass
IMP	Inosine monophosphate
IVT	In vitro transcription
kb	Kilo base pairs
KOD	Thermococcus kodakaraensis
L or Lig	Ligand
LB	Luria Broth
LP	Lateral patches
MABT	Maleic acid buffer containing Tween 20
MATLAB	Matrix Laboratory
MI	Melanoiridoblast
MIMs	Multi-input motifs
MiT	Micropthalmia/Tfe

MO	Morphant or morpholino
mRNA	Messenger ribonucleic acid
MXIG	Melano-xantho-irido-glioblast
nBLAST	Nucleotide BLAST
NC(C)	Neural crest (cell)
NCBI	National Center for Biotechnology Information
NCSCs	Neural crest stem cells
NHEJ	Non-homologous end-joining
NPB	Neural plate border
NT	Neural tube
NTP	Nucleoside triphosphate
ODE	Ordinary differential equation
PAM	Photospacer adjacent motif
PBS (T)	Phosphate buffered saline (containing Tween20)
PCR	Polymerase chain reaction
PFA	Paraformaldehyde
poly(A)	Polyadenylation
PPE	Pre-placodal ectoderm
RA	Retinoic acid
RPE	Retinal pigment epithelium
rpm	Revolutions per minute
SNP	Single nucleotide polymorphism
SOC	Super optimal broth
SSC	Saline sodium citrate
SSC	Saline sodium citrate
SSCT	Saline sodium citrate containing Tween-20
ssRNA	Single stranded RNA
TAE	Tris-Acetate-EDTA
TBE	Tris-Borate-EDTA
TILLING	Targeting Induced Local Lesions in Genomes
tNC(C)	Trunk neural crest (cell)
TRI	TRIzol
U	Units (enzyme activity)
UV-Vis	Ultraviolet-visible
WT	Wild-type
ZFIN	Zebrafish Information Network

1

Introduction

1.1. The neural crest

The neural crest (NC) is a conserved vertebrate structure arising during early stages of embryogenesis. Multipotent progenitors comprise the NC and give rise to an array of widely diverse cell types, such as neurons and glial cells, bone and cartilage derivatives, adrenal and pigment cells. Not only are the NC and its derivatives indispensable for embryonic development, but they also constitute attractive models to study the molecular basis of stem cell fate choice. This section outlines NC development and focuses on the chosen models for this study, namely the zebrafish NC and its derivative chromatophore lineages. These systems are used to elucidate the gene regulatory networks (GRNs) that guide multipotent cells to adopt distinct fates.

1.1.1. NC development

The first step towards NC formation occurs prior to gastrulation, when a complex network of morphogen gradients specifies the neural from the non-neural ectoderm in the anterior side of the embryo (Tuazon and Mullins 2015). As a consequence, the neural plate border (NPB), later giving rise to the NC and the pre-placodal ectoderm (PPE), arises during gastrulation at the boundary between neural and non-neural ectoderm (Fig. 1.1 A). As neurulation proceeds, the neural plate folds to form the neural tube, which goes on to generate the central nervous system (CNS), and the overlying non-neural ectoderm, which gives rise to the epidermis (Fig. 1.1 B-D) (Purves et al. 2012; Scott. F. Gilbert 2013). Concomitantly, complex genetic interactions guide specification of the NC from the NPB. The specified NC is located on the dorsal side of the neural tube upon its closure and comprises a transient population of multipotent cells (NCCs; Fig. 1.1 D). This structure extends dorsally from the head, through the trunk and towards the posterior-most parts of vertebrate embryos. At least a subset of NCCs have been demonstrated in culture to possess stem cell characteristics and are thus also referred to as NC stem cells (NCSCs) (Stemple and Anderson 1992).

Following induction and specification of NCCs, the next step involves acquisition of mesenchymal characteristics and of the capacity to migrate away from the neural tube. Upon closure of the neural tube, NCCs embedded within the structure undergo an epithelial-to-mesenchymal transition (EMT), by which they detach and disperse to form a loose population of cells residing dorsally and dorsolaterally to the prospective CNS (Theveneau and Mayor 2012). Delaminated NCCs acquire highly invasive properties that allow them to migrate through tissues in a spatiotemporally coordinated fashion, so that precursors of each derivative occupy the appropriate positions in the developing embryo.

In the vertebrate trunk, NC migration takes place via two main pathways: the dorso-lateral and the ventro-medial pathway. The former guides cells through the space between mesodermal tissue (dermomyotome in chick and mammals, somites in zebrafish) and the epidermis, while the latter houses migration in close proximity to the neural tube (Fig. 1.1 D). In chick and mammalian embryos, ventro-medially migrating NCCs either pass through the sclerotome, or migrate in the space between the sclerotome and the notochord, whereas in zebrafish they only pass between somite blocks, comprising both muscle and sclerotome cells (Morin-Kensicki and Eisen 1997), and the neural tube (Robert N Kelsh et al. 2009; Theveneau and Mayor 2012). Several lines of evidence suggest that NC derivatives migrate in a fate-specific manner (Erickson and Reedy 1998; R N Kelsh, Schmid, and Eisen 2000; Robert N Kelsh et al. 2009).

Specified NCCs undergo highly dynamic processes to generate the vast diversity of NC derivatives (Morriss-Kay, Ruberte, and Fukiishi 1993; Nicole M Le Douarin and Dupin 2003). Owing to the astonishing array of both ectomesenchymal and non-ectomesenchymal derivative lineages, the NC has been referred to as the ‘fourth germ layer’ (B K Hall 2000). The majority of the craniofacial cartilage and bones of the head are derived from the NC (Kague et al. 2012; Knight and Schilling 2013). Non-ectomesenchymal NC derivatives include the majority of the constituents of the peripheral nervous system (PNS) (M Bronner-Fraser and Fraser 1988; Raible et al. 1992; Bixby et al. 2002; Kruger et al. 2002) and the pigment cells found in each organism (J T Bagnara et al. 1979; M Bronner-Fraser and Fraser 1988; R I Dorsky, Moon, and Raible 1998; Bennett and Lamoreux 2003). Moreover, NCCs are progenitors of endocrine chromaffin cells (M Bronner-Fraser and Fraser 1988; Vogel 1996) and of smooth muscle cells (Hirschi and Majesky 2004; Ando et al. 2016).

Due to the aforementioned features, the NC constitutes an attractive model for studying developmental biology. Specifically, the complex and highly dynamic processes leading to the generation and maintenance of multipotent NCCs, the ability to invade tissues and achieve long-range migration, and the capacity to generate widely diverse lineages in a spatiotemporally organised manner are primarily explored features. Thus, the stem cell

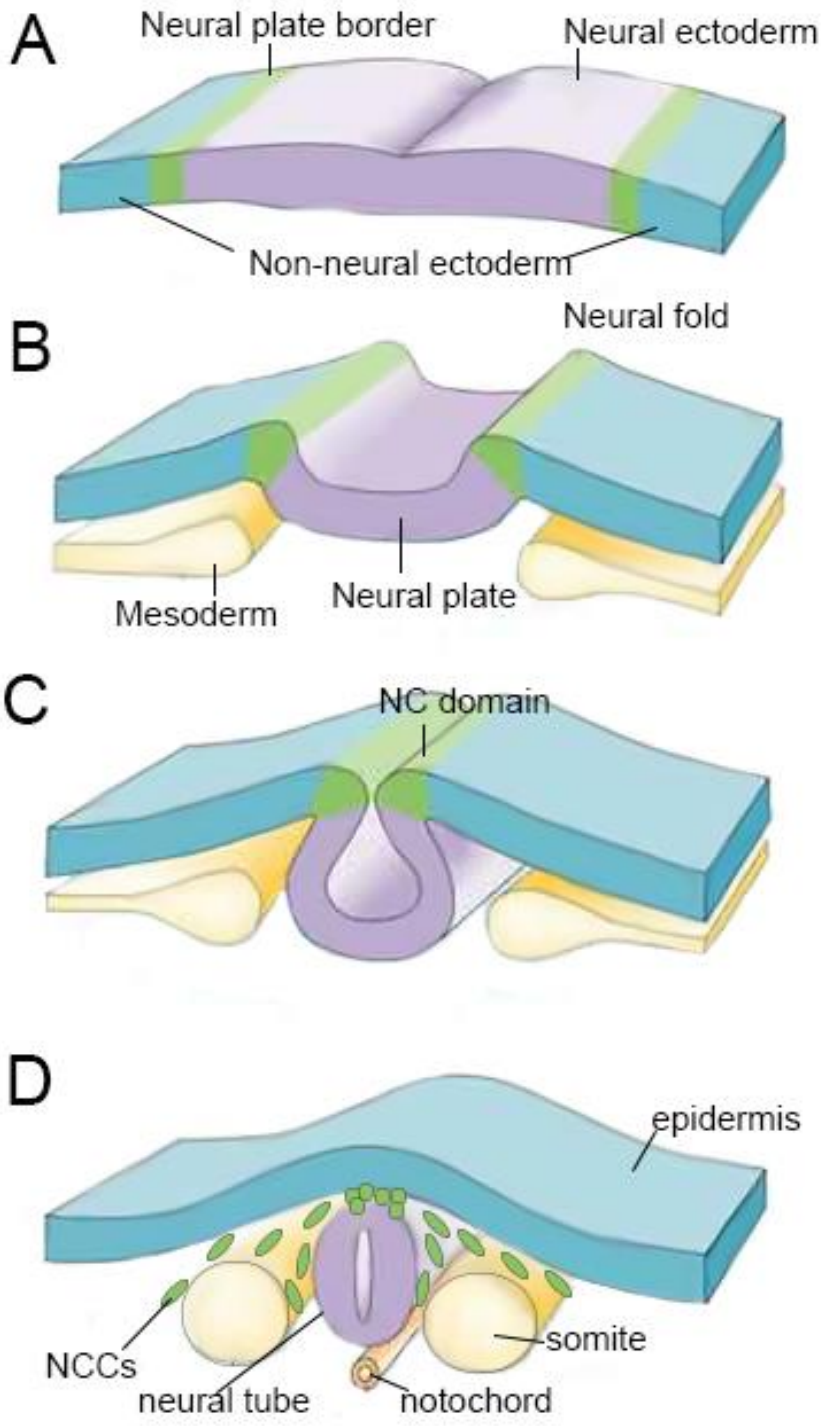


Figure 1.1. Establishment of the NC. (A) The NPB (green) arises between the neural (purple) and the non-neural (blue) ectoderm domains during gastrulation. (B, C) The NC domain (green) becomes specified from the NPB during neurulation. (D) NCCs (green circles) delaminate from the dorsal neural tube (purple) and begin to migrate either between the somites (yellow) and the epidermis (blue) or between the somites and the neural tube. Adapted from (Gammill and Bronner-Fraser 2003).

and cancer fields have greatly benefitted from elucidating the mechanisms behind NC development (Kulesa et al. 2006; Ramgolam et al. 2011; Hauser et al. 2012; Dupin and Coelho-Aguiar 2013). Although deciphering the molecular basis of NC development has been a major objective for several years, constructing the complex and highly dynamic underlying GRNs has only more recently attracted attention (Sauka-Spengler and Bronner-Fraser 2008; M. Simões-Costa et al. 2015). Although a large number of research groups have invested in constructing the GRNs that guide NCC specification from the NPB and generation of their derivative lineages (examples of experimentally determined GRNs are further discussed in section 1.3), important aspects of the process remain elusive.

1.1.2. Neural crest subtypes and derivative lineages

The premigratory NC can be classified into four functional subdomains according to location along the rostro-caudal axis of the embryo and the type of derivatives generated: the cranial NC (cNC), the cardiac NC, the vagal and sacral NC and the trunk NC (tNC) (S. F. Gilbert 2000). The following subsections describe the identified NC subdomains and introduce their respective derivatives.

1.1.2.1. Cranial NC

Over three decades ago, Gans and Northcutt proposed that vertebrate evolution depended upon development of head structures, particularly the facial compartment, and that these structures were dependent upon the NC (Gans and Northcutt 1983). In the following decades, studies revealed that precursors of the cNC migrate in distinct, organised streams (Graham, Begbie, and McGonnell 2004; Steventon, Mayor, and Streit 2014), giving rise to a wide array of constituents of the vertebrate head (Santagati and Rijli 2003; Dupin, Creuzet, and Le Douarin 2006; Minoux and Rijli 2010) .

The cNC is responsible for generating bones, cartilage and connective tissue as well as peripheral neurons and glia of the anterior head (Lumsden, Sprawson, and Graham 1991; Schilling and Kimmel 1994). Specifically, the rostral-most migratory stream of NCCs (the mandibular stream) gives rise to jawbones and a subset of the ear bones. The same stream has also been shown to generate a subset of neurons of the trigeminal ganglion. The hyoid stream gives rise to sensory facial ganglia, as well as the hyoid skeleton. The caudal-most stream, called the branchial stream, generates peripheral neurons and glia of the neck, as well as the thymus, thyroid and parathyroid glands and skeletal elements associated with these organs (Lumsden, Sprawson, and Graham 1991; Schilling and Kimmel 1994; Bockman and Kirby 1984). In addition to these streams, the fronto-nasal region is also used by rostro-ventrally migrating cNCCs which

proceed to generate facial bones and cartilage (S. F. Gilbert 2000; Steventon, Mayor, and Streit 2014).

Furthermore, the cNC has been shown to be a crucial contributor in development of dental tissues and to give rise to dental pulp stem cells, which are capable of differentiating towards several NC derivatives *in vitro* (Chai et al. 2000; Janebodin et al. 2011). Embryonic melanophores, one of the three pigment cell types in zebrafish, have also been shown to arise from cNCCs (Schilling and Kimmel 1994). *In vitro* culture of chick NCCs suggested the existence of common progenitors between pigment cells, cartilage and neural derivatives (Baroffio, Dupin, and Le Douarin 1991). In zebrafish embryos, neurons were not found to share progenitors with cartilage cells. Instead, NCCs capable of generating pigment, cartilage and glial derivatives were identified (Schilling and Kimmel 1994). The melanophore lineage will be discussed in more detail in sections 1.2.2.4 and 1.2.3.1.

Other than generating the aforementioned derivatives, the cNC has been reported to play important roles in patterning the bones and muscles of the vertebrate head, as a result of instructive cues likely including BMP and Wnt signalling (Marcucio et al. 2011; Rinon et al. 2007). Moreover, the developing brain receives vital signals from the cNC during embryogenesis (Creuzet, Martinez, and Le Douarin 2006).

1.1.2.2. Cardiac NC

The cardiac NC has been identified in the early 1980s in avian embryos (Kirby, Gale, and Stewart 1983; Stoller and Epstein 2005). Specifically, it was shown that cardiac NC progenitors arise caudally to the cNC, between the mid-otic placode and the third somite pair (Kirby, Gale, and Stewart 1983) and contribute towards the formation of the heart and the major vessels; the aortico-pulmonary septation complex, the tunica media of the aortic arch, and its major branches are all derived from cardiac NCCs (Stoller and Epstein 2005). In zebrafish, progenitors originated from axial levels rostral to the first somite pair were shown to make a functionally significant contribution to both the embryonic heart and major vessels (Y.-X. Li et al. 2003).

1.1.2.3. Vagal NC and sacral NC

Across vertebrates, the vagal NC and sacral NC contain the only embryonic progenitors able to generate enteric neurons and glia (N M Le Douarin and Teillet 1973; R N Kelsh and Eisen 2000; X. Wang et al. 2011). In chick and mouse embryos, the vagal NC extends from the 1st to the 7th pair of somites, thus partly overlapping with the cNC segment, whereas the sacral NC is located posterior to the 28th somite pair in chick (N M Le Douarin and Teillet 1973; Epstein et al. 1994; Burns, Delalande, and Le Douarin 2002) and to the 25th in mouse (Heanue, Shepherd, and Burns 2016). In zebrafish, sacral

NC has not been experimentally identified, thus rendering the vagal NC responsible for generating the entire population of enteric neurons and glia (Olden et al. 2008; Shepherd and Eisen 2011; Heanue, Shepherd, and Burns 2016). NC progenitors fated to occupy the developing gut first migrate towards the ventral side of the embryo, from where they enter the foregut (or the intestinal bulb of the zebrafish embryo) and migrate caudally to colonise the remaining gut (Heanue, Shepherd, and Burns 2016). In avian and mammalian embryos, the sacral NCCs also migrate ventrally and proceed to enter the hindgut. It has been suggested that sacral NC derivatives colonise a small portion of the gut by migrating along nerve fibers derived from pelvic ganglia (X. Wang et al. 2011).

1.1.2.4. Trunk NC

The trunk NC (tNC) extends between the vagal and the sacral NC (somite pairs 8-28 in chick). In zebrafish embryos, the tNC is considered to overlap with the vagal and partly with the cardiac NC, as it is associated with somite pair 1, and to extend to the tail of the embryo (Kimmel et al. 1995). In the majority of the literature, no distinction is made between the NCCs occupying the trunk region of the embryo and those of the tail region. Instead both of these domains are considered as tNC. Nevertheless, in the context of mesoderm induction, the trunk and tail regions are regarded as separate domains, subject to distinct regulatory inputs (Kimelman 2006). It is therefore meaningful to explicitly investigate the differences between anterior (trunk) and posterior (tail) tNC.

The tNC generates neuronal, glial, endocrine and pigment cell derivatives (M Bronner-Fraser and Fraser 1988; Marianne Bronner-Fraser and Fraser 1989). In all vertebrate models studied, tNC derivatives migrate from the dorsal side of the neural tube towards the ventral side of the embryo using both the dorso-lateral and ventro-medial pathways. In the zebrafish context, these routes are often referred to as lateral and medial pathways, respectively. In zebrafish embryos, PNS constituents start migrating prior to other derivatives through the medial pathway. Contrary to other models, pigment precursors in zebrafish follow the same route, doing so approximately 2 hours after the start of migration (Robert N Kelsh et al. 2009). The dorso-lateral pathway becomes occupied by additional pigment precursors with a further 2 to 3 hour delay (Robert N Kelsh et al. 2009).

All of the glial and a significant proportion of the neuronal constituents of the trunk PNS are derived from the tNC (N. Le Douarin and Kalcheim 1999). In mouse, chick, *Xenopus* and zebrafish embryos, tNCCs giving rise to components of the dorsal root ganglia (DRG) migrate ventro-medially, along the rostral half of each somite. Upon reaching their respective positions on either side of the spinal cord, gradual differentiation of sensory neuron subtypes takes place. These cells extend their axons to innervate the skin,

tendons and muscles (Raible and Ungos 2006; Prendergast and Raible 2014). Furthermore, ventro-medially migrating tNCCs in all model species give rise to glial cells. Satellite glia are associated with neuronal cell bodies within the DRGs, whereas myelinating and non-myelinating Schwann cells are associated with nerve axons (Woodhoo and Sommer 2008).

Migrating ventro-medially along with sensory neuron progenitors are cells that will later give rise to the sympathetic ganglia, located adjacent to the dorsal aorta (Raible and Ungos 2006; An, Luo, and Henion 2002). Sympathetic ganglia predominantly consist of adrenergic neurons differentiating after the DRGs and play vital roles in regulating several of the unconscious functions of the organism. For example, they are responsible for stimulating smooth muscle cells, such as the cardiac muscle cells (Stewart et al. 2010). *In vivo* studies have shown that sympathetic neurons share a common NC-derived progenitor with chromaffin cells of the adrenal medulla, or the interrenal gland in zebrafish (Huber 2006; Y.-W. Liu 2007; Saito et al. 2012).

Finally, tNCCs give rise to the pigment cells lineages of all vertebrates. Pigment cells, also called chromatophores, play important roles during vertebrate evolution as they protect from environmental dangers, for instance by absorbing UV radiation, and providing thermoregulation (Hegna et al. 2013). Furthermore, they provide camouflage, which allows for predation and for protection from predators. Finally, in several species communication and mating choice relies upon recognition and alteration of pigment patterns. In chick and mammals there is only a single pigment cell type: the melanocyte (Mort, Jackson, and Patton 2015).

Melanocytes are highly dendritic cells containing cytoplasmic organelles called melanosomes, inside which the macromolecule melanin is synthesised. In mammals, two different types of melanin have been identified: black or dark brown eumelanin and red or yellow pheomelanin (Hearing 2011). Melanin production, a process confined to melanosomes due to its associated toxicity (Mort, Jackson, and Patton 2015), requires the enzyme tyrosinase (Tyr), which catalyses the conversion of tyrosine to DOPAquinone (Hearing 2011). To produce eumelanin, the crucial melanogenic enzymes domachrome tautomerase (Dct), tyrosinase-related protein 1 (Tyrp1) and Tyr catalyse a chain of reactions involving the intermediate molecules DOPAchrome, DOPA, 5,6-dihydroxyindole (DHI) and DHI-2-car-boxylic acid (DHICA). Pheomelanin synthesis relies upon reaction of cysteine's sulfhydryl groups with DOPAquinone and sequential modifications of resulting derivatives (Riley 1999; Hearing 2011). In humans, mutations in TYR and TYRP1 result in OCA1 and OCA3 type of albinism (Oetting 2000; Montoliu et al. 2014). Mammalian melanocytes maintain only a proportion of their melanosomes

and export the remainder to overlying keratinocytes through the melanocytic dendrites, thus protecting the skin from UV radiation (Boissy 2003).

Additional tNC-derived pigment cells have been discovered in other vertebrate species. These include light-reflecting iridophores and yellow xanthophores commonly found in fish, reptiles and amphibians. Iridophores owe their reflective properties to crystallised purine structures present within cytoplasmic organelles called reflecting platelets (Menter et al. 1979; Cloney and Brocco 1983; Schartl et al. 2016). Xanthophores contain carotenoid vesicles but primarily rely on compounds called pteridines for their pigmentation (Obika 1993). These two pigment cells will be discussed further in the zebrafish context (refer to section 1.1.3). Primarily in fish, additional types of pigment cells have been identified. These include erythrophores, which may appear orange, red or violet depending on different combinations of carotenoid vesicles with light-reflecting iridosomes, or with pteridine compounds within the cytoplasm (Matsumoto 1965; Goda et al. 2011). Leucophores are white, reflective cells owing to the presence of cytoplasmic reflective platelets (Cloney and Brocco 1983). Finally, the greenish-blue colour of cyanophores results from a molecule that has not been identified to date (Goda and Fujii 1995). It has yet to be demonstrated whether this cell type derives from the NC (Robert N Kelsh 2004).

Owing to their inherent tractability, accessibility and amenability to genetic manipulations, since they are dispensable for the survival of model organisms, chromatophores have attracted attention from geneticists and developmental biologists for over 100 years (Bennett and Lamoreux 2003; Mort, Jackson, and Patton 2015). Even though melanocytes have been in the research spotlight because they are ubiquitous in vertebrates and understanding their development has direct implications in melanoma research, the attractive properties of pigment cells are shared by the other less well-studied lineages. In our research group, the three chromatophore lineages of the zebrafish: melanophores, iridophores and xanthophores, are employed to elucidate mechanisms of NC development, in particular the molecular basis of NCC fate choice, the existence of partially restricted progenitors and of NC-derived adult stem cells.

1.1.3. Zebrafish chromatophore lineages

1.1.3.1. Melanophores

In zebrafish larvae, mature melanophores form four distinct stripes along the body: a dorsal band that extends from the dorsal head to the tail, one lateral stripe on each side of the embryo, positioned along the horizontal myoseptum, a ventral band stretching from the head towards the tail along the dorsal side of the yolk sac and a yolk sac stripe,

visible along the ventral side of the yolk sac (Fig. 1.2 A) (Quigley and Parichy 2002; Robert N Kelsh et al. 2009).

Melanophores are the zebrafish equivalent to mammalian melanocytes and share several of their properties in terms of morphology and transcriptional regulation. Like their mammalian counterparts, melanophores are highly dendritic cells (Fig. 1.2 B) that owe their black pigmentation to eumelanin, synthesised within melanosomes. Unlike mammalian melanocytes, melanophores maintain all of their melanosomes within their cytoplasm (Mort, Jackson, and Patton 2015).

The precursor of the melanophore is called the melanoblast. Melanoblasts become specified from NC progenitors relatively early, likely prior to the onset of migration. This can be concluded based on early expression of melanophore markers, such as *dct*, and on the fact that melanoblasts migrate over narrow time windows through specified pathways, not only in the zebrafish but also in other model organisms (Serbedzija, Fraser, and Bronner-Fraser 1990; Robert N Kelsh et al. 2009; Theveneau and Mayor 2012). In zebrafish, approximately 2 hours following the onset of migration, specified melanoblasts follow the ventro-medial pathway. Between 2 to 3 hours later, they enter the dorsolateral pathway. For simplicity, these two pathways will from now on be referred to as the medial and lateral pathways, respectively. In zebrafish, not only specified but also differentiating melanocytes, identified based on deposition of melanin pigment, are found along migratory pathways from approximately 26 hpf, providing further support to the notion that lineage specification occurs prior to NCCs reaching their final locations along the embryonic body.

1.1.3.2. Iridophores

In zebrafish, mature iridophores appear as silvery/white spots and clusters under incident light (Fig. 1.2 E; Fig. 1.3) (Robert N Kelsh 2004). Embryonic iridophores first become apparent in the eye at 42 hpf, while iridophores of the trunk are clearly distinguished by 48 hpf, when they occupy the different stripes in association with melanophores (Kimmel et al. 1995). Dense clusters of iridophores appear in the developing lateral patches (LPs) and arranged on either side of the developing swim bladder, above the yolk sac (Fig. 1.2 E; Fig. 1.3) (Kimmel et al. 1995). Finally, iridophores are spread across the eye, positioned peripheral to the retinal pigment epithelium (RPE) (Fig. 1.3).

Iridophores generate the characteristic silver colouration of the body owing to the presence of crystallised purines within cytoplasmic reflecting platelets, also known as iridosomes (Fig. 1.2 D; Fig. 1.3). However, the structures formed do not correspond to ‘true pigments’, such as melanin and carotenoids, they instead generate structural colouration (J T Bagnara et al. 1979). Structural pigmentation (Fox 1976; Land 1972)

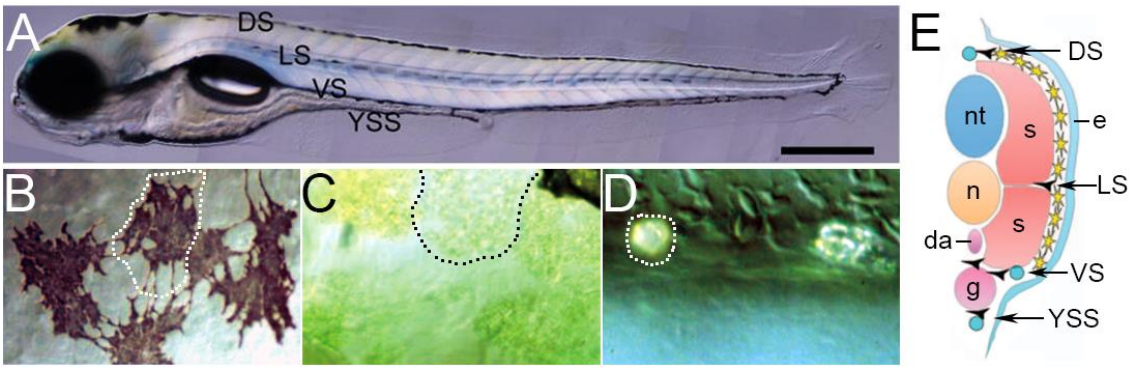


Figure 1.2. Zebrafish pigment cell types and distribution. (A) Larval pigment cell pattern. Melanophores occupy the dorsal (DS), lateral (LS), ventral (VS) and yolk sac (YSS) stripes, as well as the eye of the larva. Adapted from Kelsh et al. 2009. (B-D) High magnification views of (B) melanophores, (C) xanthophores and (D) iridophores. Single cells are outlined. Adapted from Kelsh 2004. (E) Schematic of a cross-section across the larval trunk (right half only). Locations of melanophores (black shapes), iridophores (light blue circles) and xanthophores (yellow shapes) in relation to the neural tube (nt), notochord (n), dorsal aorta (da), gut (g) and somites (s) are represented. Adapted from Kelsh 2004.

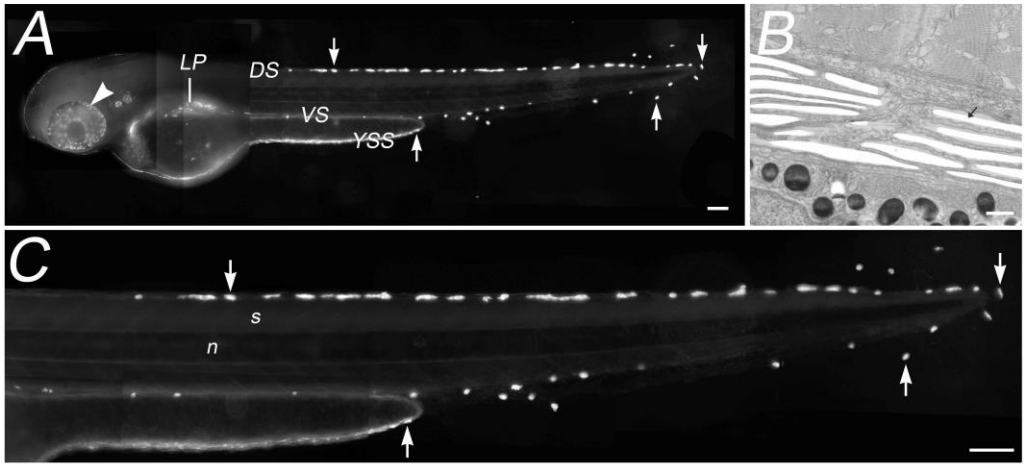
has been shown to generate the astounding colours of an array of vertebrate classes, including fish, reptiles, birds and mammals (Joseph T Bagnara, Fernandez, and Fujii 2007), as well as of invertebrates, such as cephalopods (Andouche et al. 2013). It owes itself to an optical phenomenon known as constructive interference (Land 1972; Joseph T Bagnara, Fernandez, and Fujii 2007), which occurs when light is reflected by alternating layers of materials (in the case of iridosomes layers of purine crystals separated by gaps; Krauss et al. 2013) with high and low refractive indices, which are of specific thickness compared to the light wavelength (Land 1972). Light waves scattered from distinct alternating layers are 'in phase' and, therefore, enhance each other.

The major purine constituents are guanine, adenine, hypoxanthine and uric acid (J T Bagnara et al. 1979). Studies on the metabolic pathways involved in purine biosynthesis (Fig. 1.3 D), complemented by investigation of the iridophore transcriptome, have highlighted a complex network of enzymatic reactions (Moriwaki, Yamamoto, and Higashino 1999; Higdon, Mitra, and Johnson 2013). Key facilitators for guanine and adenine biosynthesis are the enzymes purine 5'-nucleotidase and purine nucleoside phosphorylase (PNP). The former dephosphorylates guanosine and adenosine monophosphate (GMP, AMP) to generate guanosine and adenosine, respectively. Then, PNP cleaves guanosine and adenosine to produce guanine and adenine, respectively. Hypoxanthine can be synthesised through cleavage of inosine by PNP. Inosine is itself generated by degradation of adenosine through the enzyme adenosine deaminase (ADA). Finally, uric acid is produced by subsequent oxidisation reactions of hypoxanthine to xanthine and of xanthine to uric acid, mediated by the enzyme xanthine oxidase (XO).

The NC origin of iridophores has been proven using lineage tracing studies (Dutton et al. 2001), as well as transgenic lines which permanently label NC derivatives (Rodrigues et al. 2012; Mongera et al. 2013). Moreover, complete lack of iridophores has been demonstrated upon loss of function of genes important for NC development, in particular *sox10* (R N Kelsh et al. 1996). Iridoblasts, the specified precursors of iridophores marked by expression of the receptor tyrosine kinase (RTK), leukocyte tyrosine kinase (*ltk)*, are present from early stages of NC development (Lopes et al. 2008). They migrate exclusively through the medial pathway and populate the dorsal, ventral and yolk sac stripes where they differentiate into mature cell types.

1.1.3.3. Xanthophores

Yellow xanthophores (Fig. 1.2 C) are pigmented by pteridines, contained within organelles called pterinosomes, as well as carotenoid compounds (Obika 1993; Ziegler 2003). Important enzymes catalysing reactions within the intricate pteridine biosynthesis



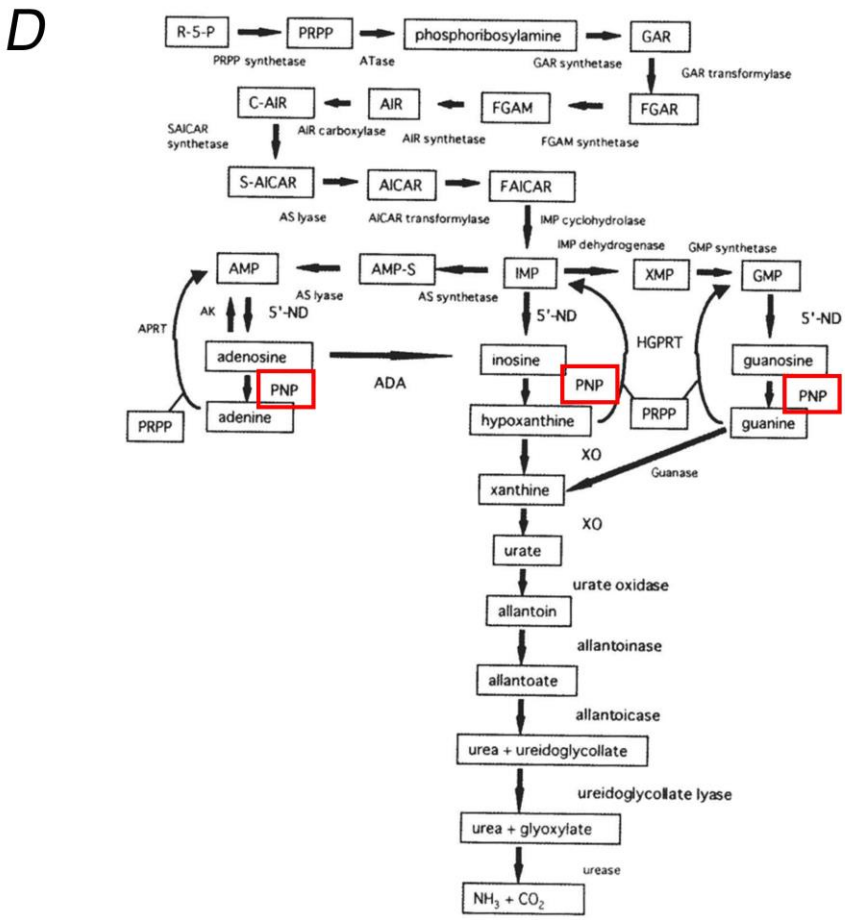


Figure 1.3. Distribution of embryonic iridophores at 2 dpf. (A, C) Differentiated iridophores (arrows) viewed under incident light along the dorsal (DS), ventral (VS) and yolk sac (YSS) stripes, as well as overlaying the RPE (arrowhead) and on the lateral patches (LPs). Embryos oriented laterally with the head towards the left. (B) Electron microscopy section showing iridosome morphology and distribution within the iridophore cytoplasm. Adapted from Krauss et al. 2013. (D) The metabolic pathways of purine biosynthesis. The key enzyme PNP, which is one of the candidate genes for the iridophore GRN, is highlighted. Adapted from Moriwaki et al. 1999. Scale bars: (A, C) 100 μm; (B) 0.5 μm.

pathway include GTP cyclohydrolase (*gch*), xanthine oxidase (*xod*) and xanthine dehydrogenase (*xdh*). Genes coding for these enzymes are expressed significantly prior to xanthophore differentiation along the premigratory NC domain and within the migratory pathways of the developing zebrafish embryo (Ziegler 2003). Xanthophore precursors, called xanthoblasts, migrate from the tNC strictly through the lateral pathway at the same time as melanoblasts (Robert N Kelsh 2004) and gradually occupy the intervening spaces between iridophores and melanophores (Fig. 1.2 C) (Odenthal, Rossnagel, et al. 1996; Quigley and Parichy 2002; Robert N Kelsh et al. 2009). Differentiated xanthophores are visible from approximately 3 dpf.

1.1.4. The direct and progressive fate restriction models

The mechanisms and processes underlying fate choice of NCCs towards so many different cell types have long been the subject of investigations. There are two currently supported models aiming to describe NC specification, the so-called direct fate restriction model and the progressive fate restriction model. The direct fate restriction model has been given a lot of support (Marianne Bronner-Fraser and Fraser 1989, 1991; Collazo, Bronner-Fraser, and Fraser 1993; Delfino-Machin et al. 2007), arguing that NCCs migrate and become patterned while maintaining full multipotency before differentiating to each cell type in a manner dependent upon local environmental cues (R. Kelsh and Erickson 2013).

An alternative model that describes diversification from the NC is the progressive fate restriction model, under which premigratory NCCs become partially specified prior to reaching their final destination, and perhaps while they are still located dorsally to the neural tube (Weston 1991; Henion and Weston 1997; Nicole M Le Douarin, Calloni, and Dupin 2008). This model proposes that sequential partial fate restriction events ultimately give rise to unipotent precursors, such as melanoblasts, iridoblasts and xanthoblasts, which ultimately differentiate to form mature NC derivatives.

In vivo evidence in support of the progressive fate restriction model stems from lineage-tracing studies, in which premigratory NCCs were labelled either with a vital dye, or with a recombinant retrovirus (M Bronner-Fraser and Fraser 1988; Frank and Sanes 1991). Although these studies demonstrated the existence of NCCs giving rise to diverse cell types, they also revealed that a significant proportion of cells located dorsally to the neural tube could only generate a limited number of derivatives, or even give rise to homogeneous clones. Therefore, these results indicated that a proportion of NCCs become fate-restricted prior to emigration from the neural tube. Moreover, it has been suggested that, following delamination and prior to the onset of migration, the arrangement of premigratory cNCCs dorsolaterally of the neural tube in zebrafish tightly

correlates with the fate that each cell is biased towards. Specifically, laterally positioned cells gave rise predominantly to neuronal lineages, medially positioned cells were biased towards chromatophore and cartilage fates, whereas progenitors located closest to the neural tube generated cartilage and connective tissue constituents (Schilling and Kimmel 1994). An additional line of evidence in support of the progressive fate restriction model came from elegant lineage tracing experiments carried out in chick embryos (Krispin et al. 2010). In this study, it was demonstrated that NCCs were fate restricted prior to emigration from the neural tube and that, at least in the case of neuronal precursors, exposure to different microenvironments *in vivo* was unable to alter the chosen pathway of these cells.

Based on the aforementioned evidence, it is conceivable that truly multipotent NCCs exist over a restricted window of time, with cell lineages segregating surprisingly early and even prior to the onset of migration. Lineage-tracing studies both *in vitro* and *in vivo* (Baroffio, Dupin, and Le Douarin 1988; M Bronner-Fraser and Fraser 1988; Frank and Sanes 1991) frequently identified clones deriving from single NCCs that did not generate the entire complement of NC derivatives. Instead, two or three different cell types were generated, suggesting that partial fate segregation occurs even prior to emigration from the neural tube. In summary, compelling evidence supports the validity of the progressive fate restriction model, but it is important to note that no definitive conclusions can be drawn with regard to the potential of individual NCCs under physiological conditions and across species. This is because the discussed experiments were conducted both *in vitro* and *in vivo*, using different models and conditions.

By definition, it follows that intermediate progenitors of different derivatives would be present during specification stages. Indeed, investigations to date have directly identified the presence of bipotent or oligopotent progenitors. For example, it has been shown in chick embryos that sympathetic derivatives of the PNS and chromaffin cells of the adrenal medulla share a common sympathoadrenal progenitor, development of which, as well as further fate restriction, are processes dependent upon BMP signalling from the dorsal aorta (Huber 2006; Saito et al. 2012). In both mouse and avian embryos, a common progenitor between melanocytes and glial cells has been identified, which is referred to as the melanoglioblast (Thomas and Erickson 2009; Adameyko et al. 2009; Adameyko and Lallemend 2010). Moreover, in zebrafish, a common progenitor of neurons and glial cells has been demonstrated (McGraw, Nechiporuk, and Raible 2008). Additionally, lineage tracing studies on zebrafish embryos demonstrated the existence of a common progenitor capable of generating both melanophores and iridophores, termed a melanoiridoblast (Curran et al. 2010).

Previous work from our group provided evidence for the existence of an early common progenitor between all chromatophore lineages, and possibly also of glial cells, the chromatoglioblast. The presence of a common progenitor for all chromatophore lineages (chromatoblast) was first proposed in the late 70s, when mosaic cells containing more than one type of pigment organelle were described (J T Bagnara et al. 1979). Results from phenotypic analyses of zebrafish *sox10* mutants were consistent with this concept (Dutton et al. 2001). Upon loss of *sox10* function, defects in the development of diverse NC derivatives, including neurons, glia and pigment cells, were observed. Interestingly, although migration of neuronal derivatives towards the ventral side of the embryo appeared grossly unaffected in mutant embryos, pigment cell lineages failed to become specified (Dutton et al. 2001). Importantly, *sox10* mutant embryos presented with cells trapped within the premigratory NC domain of the dorsal trunk over prolonged periods of time, while maintaining expression of the iridophore marker, *ltk*. The numbers of these cells were considerably higher than the estimated number of iridophore progenitors expressing *ltk* in WT embryos and they were shown not to maintain expression of other multipotent NC markers, thus suggesting that they represented partially fate-restricted non-neuronal progenitors (Lopes et al. 2008). Subsequent studies using an *ltk* reporter construct suggested that both chromatophore and glial lineages are derived from *ltk* positive precursors (Nikaido, M., unpublished data).

Consistent with the data on partially restricted progenitors, the working model for iridophore development adopted in this thesis is shown in figure 1.4. Expression of candidate genes for the iridophore GRN was first evaluated at 18 hpf within the premigratory NC domain, where the chromatoglioblast (MXIG) was postulated to reside. Published data to date suggested that the MXIG is positive for *sox10* and *ltk* expression (Dutton *et al.*, 2001; Lopes *et al.*, 2008; M. Nikaido, unpublished data), but negative for *mitfa* expression. Expression of the latter gene has been experimentally detected only in very few dorsally located cells at 18 hpf (J A Lister et al. 1999). Subsequent experimental time points were selected to identify gene expression during consecutive transitions of the MXIG to the melanoiridoblast (MI), the iridoblast (I) and, ultimately, the differentiated iridophore. Based on the *sox10* expression pattern as it has been thus far established (Dutton et al. 2001; Takada and Appel 2011; Greenhill et al. 2011; Kwak et al. 2013), and considering previously conducted studies directly testing presence of *sox10* transcript and protein in developing melanocytes (Greenhill et al. 2011), it is concluded that *sox10* is likely active at the MI stage and becomes gradually downregulated upon establishment of pigment cell fates. Furthermore, it has been reported that MI fate choice depends upon activation of *foxd3* in a subset of progenitors, resulting in repression of *mitfa* and, thus, in the establishment of the iridoblast (Curran et al. 2010).

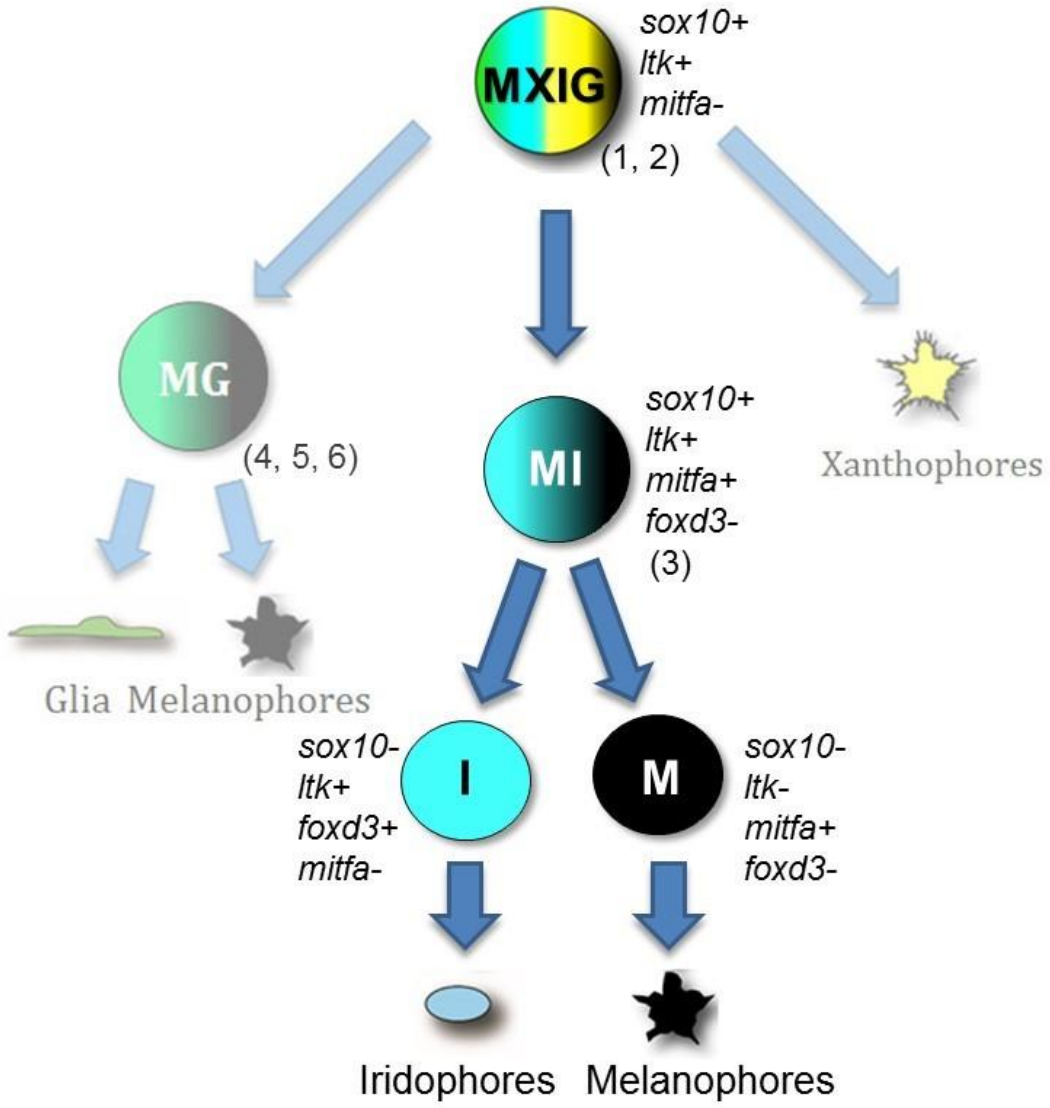


Figure 1.4. The working model for iridophore development. According to previous data, the tetrapotent chromatoglioblast (or melano-xantho-irido-glioblast, MXIG) (**1**, Lopes et al. 2008; **2**, M. Nikaido, unpublished data) generates the bipotent melanoiridoblast (MI) (**3**, Curran et al. 2010), which, following *foxd3*-dependent *mitfa* repression (**3**) gives rise to either an iridoblast (I), or to a melanoblast (M) if *mitfa* remains activated. These unipotent cells proceed to differentiate to mature cell types. Bipotent precursors generating xanthophores have not been observed. Another bipotent progenitor, the melanoglioblast, has been demonstrated in mouse (**4**, Adameyko et al. 2009; **5**, Adameyko & Lallemend 2010) and chick (**6**, Thomas & Erickson 2009) embryos. Key genes suggested to function in precursors of the iridophore and melanophore lineages are indicated.

1.2. Gene regulatory networks in development

1.2.1. Developmental GRNs

During embryogenesis, highly dynamic cellular events occur in concert, ultimately leading to the formation of complex organisms. These events begin from the earliest stages of fertilisation and include coordinated proliferation, timely patterning, programmed cell death and diversification of distinct cell types (Plickert, Kroiher, and Munck 1988; Abrams et al. 1993; Zernicka-Goetz 2002; C. Y. Leung and Zernicka-Goetz 2015). The spatiotemporal coordination of these processes is under the control of tightly regulated transcriptional programmes, triggered at different times and within distinct cell types (Levine and Davidson 2005). Activation of transcriptional programmes is dependent upon the action of context-specific transcription factors (TFs), which are proteins able to bind particular regions of DNA. TFs associate with cis-regulatory elements of their target genes to activate or inhibit their expression (Spitz and Furlong 2012). Such regulatory elements include promoters, enhancers, silencers or insulators, and they are typically bound by several TFs simultaneously, a feature which fine-tunes transcriptional regulation in space and time. Activated genes encode for other TFs, but also transmembrane receptors, epigenetic modulators, enzymes or non-coding regulatory elements such as microRNAs, which harmoniously facilitate a particular cellular event, or series of events.

Gene regulatory networks (GRNs) are maps aiming to logically and in detail describe the regulatory interactions between sets of genes, which cooperatively determine cellular phenotypes (Levine and Davidson 2005). GRN complexity ranges vastly from very small to large and extremely complex networks (Zhou et al. 2007; Krumsiek et al. 2011; Kueh and Rothenberg 2013). During development, specialised and highly dynamic GRNs have been described, which dictate proliferation, patterning, differentiation, inter-cellular communication and migration in response to extracellular signals (Theveneau and Mayor 2012; Balaskas et al. 2012; M. Simões-Costa et al. 2015). Moreover, the establishment, maintenance and fate segregation of stem cells constitute aspects of development controlled by intricate GRNs. Stem cells are characterised by two major properties: firstly, they can maintain their populations, or give rise to clones of themselves through cell division and, secondly, they can become specified and differentiate into diverse lineages (Nadig 2009). Both properties are greatly dependent upon GRNs functioning within these cells to dynamically control gene expression and, ultimately, the fate of a stem cell (Boiani and Schöler 2005; Chan, Yang, and Ng 2011).

An emerging model to capture the GRNs underlying stem cell dynamics *in vivo* is the neural crest (NC). The NC is a common feature in vertebrate embryogenesis and comprise a transient population of multipotent progenitors, capable of giving rise to

widely diverse cell types and, indeed, to stem cells that persist into adulthood (Shakhova and Sommer 2008). Due to the exciting features of the NC, which are detailed in section 1.1, and the stem cell-like properties of NC progenitor cells, the structure has attracted attention over several decades (Gans and Northcutt 1983; N. Le Douarin and Kalcheim 1999; B K Hall 2000; B. K. Hall 2008). More recently, investigations focused on assembling the GRNs underlying the different stages of NC development (Bhat, Kwon, and Riley 2013; M. Simões-Costa et al. 2015). Despite extensive efforts, several aspects of these complex GRNs remain unclear.

1.2.2. GRNs comprise different structural motifs

From the smallest to the largest GRNs, common and recurring structural motifs can be distinguished, only some of which are listed in this section (Fig. 1.5) (Swiers, Patient, and Loose 2006; Alon 2007; Burda et al. 2011; Murugan 2012). In some cases, formation of positive or negative self-regulatory loops is observed. The former are evident when gene X is required to sustain its own expression even in the absence of its inducer. Positive self-regulation is associated with slow-rate responses, because production of the involved gene only peaks following the concentration of its product reaching its promoter's activation threshold (Alon 2007). On the other hand, negative self-regulation constitutes an efficient way to create a short burst of gene expression before a certain equilibrium, defined by the promoter's repression threshold, is reached. Negative self-regulation is common when strong promoters are involved, leading to very rapid response times in terms of generating gene product following stimulation. Upon reaching the promoter's repression threshold, production of the gene ceases and a steady-state concentration equilibrium just below this threshold is maintained (Alon 2007).

More elaborate GRN motifs include feed-forward loops (FFLs) (Alon 2007; Krumsiek et al. 2011; Murugan 2012) and multi-input motifs (MIMs) (Alon 2007). FFLs are loops comprised of three TFs X, Y and Z, where X regulates both Y and Z, and Y in turn regulates Z. These regulators could be either positive or negative, giving rise in total to eight possible loops, distinguished as either coherent or incoherent FFLs (Fig. 1.6). Each loop produces a distinct net result regarding the rates of gene expression change over time and the final equilibriums reached by the system (Alon 2007; Murugan 2012). An interesting investigation of the transcriptional response to coordinated positive and negative FFLs was published by Locke et al. (Locke et al. 2011). MIMs refer to the integration of inputs from multiple regulators on a set of genes.

1.2.3. Functional motifs in developmental GRNs

Motifs such as those described in section 1.2.2 are encountered in developmental GRNs. However, this subset of biological networks differs from others in its ability to induce

irreversible fate decisions, sustainable even after a transient developmental input seizes to exist (Alon 2007; Krumsiek et al. 2011). In other words, developmental GRNs are often capable of providing memory of an input signal, an effect which in certain cases is referred to as hysteresis (Kopfova 2006; Balaskas et al. 2012). The irreversibility or persistence of certain responses are effects mediated by FFLs and gene self-regulation, whereas distinct GRN motifs regulate fate plasticity (Alon 2007; Holmberg and Perlmann 2012). For instance, in Figure 1.5 B.a, each of the two genes (represented by nodes) is positively self-regulated and, additionally, the two genes form a double-negative feedback loop. This motif is associated with plasticity in a stem cell context, when neither gene is preferentially expressed. However, a transiently expressed upstream regulator or external cue could tip the balance and the system would lock into a steady state where one gene is on and the other is off (Crespo and del Sol 2013).

Overall, GRNs are of fundamental importance in biological and, in particular, developmental processes. Thus, the entire genome can be regarded as the components of a single GRN, with development occurring as subsets of this GRN sequentially unfold to cause progressive specification and differentiation of cell types. It follows that acquiring a deep understanding of the genetic interaction motifs and spatio-temporal dynamics of the GRNs underlying developmental processes is vital for improving our perception of embryogenesis, as well as for dissecting the extent to which random variation (stochasticity) is responsible for modulating stem cell fate choice or other biological processes.

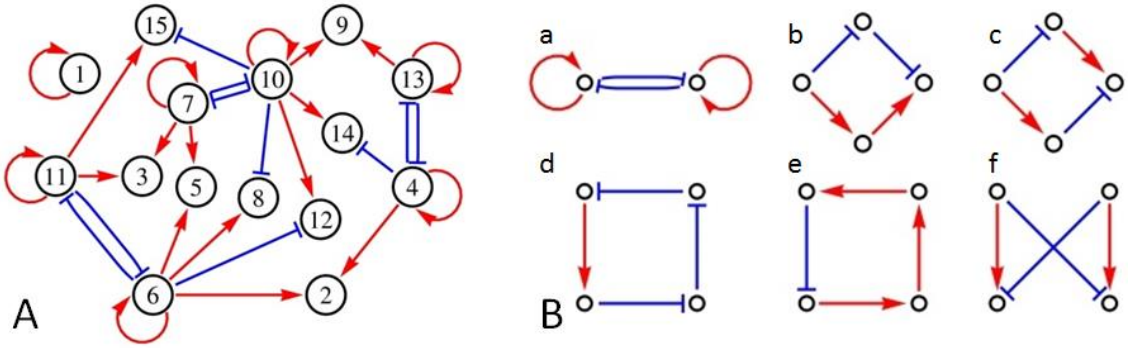


Figure 1.5. GRN structure and motifs. (A) Schematic showing the structural elements taking part in GRN formation. Genes are represented as numbered nodes. Interactions leading to transcriptional activation or repression are represented by red or blue coloured edges, respectively. The network demonstrates positive auto-regulatory effects (designated by red curved arrows, for example on node 1), simple and double negative inhibitory interactions (blue lines) as well as MIMs (red arrows from multiple nodes pointing towards one node). (B) Diagramatic presentation of distinct GRN motifs: (B.a) Double negative feedback loop with self-regulation. (B.b, B.c) Incoherent diamond. (B.d, B.e) Frustrated four-node loop. (B.f) Incoherent bifan. Adapted from (Burda et al. 2011).

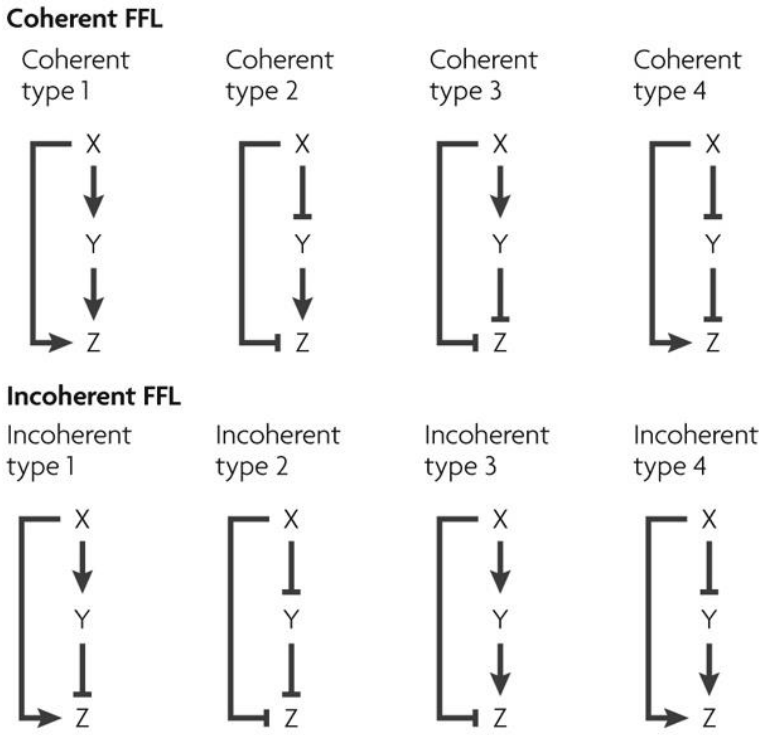


Figure 1.6. Structural representation of coherent and incoherent FFLs. FFLs are categorised as coherent or incoherent according to the nature (activating or inhibitory) and combinations of interactions that take place. Different combinations lead to either solid enhancement or suppression of Z (coherent) or to both effects taking place simultaneously (incoherent). Adapted from (Alon 2007).

1.3. The NC as a model to study GRNs

The complex GRNs underlying NC development, namely induction from the NPB, NCC specification, delamination from the neural tube, migration and commitment to different lineages have become a major research focus (Groves and LaBonne 2014; M. Simões-Costa et al. 2015). This section summarises the current understanding of interactions established between genes and signalling pathways during NC development, based on studies conducted using chick, frog, zebrafish and mouse models. As outlined in previous sections, NC development displays a high degree of conservation across vertebrate species. Therefore, although species-specific features could exist, compiling data generated using different model organisms is considered appropriate (M. Simões-Costa et al. 2015). As an example of a species-specific feature, Pax3 and Pax7 recognise similar target sequences but their NC expression patterns differ between model organisms (M. S. Simões-Costa et al. 2012). It is likely that Pax3 and Pax7 function redundantly and, thus, they have not been considered as separate members of the GRN (Simões-Costa et al. 2015; Fig. 1.8)

Intricate GRNs function during each stage of NC development, from induction of the NPB until the differentiation of diverse lineages. In the interest of space, only data concerning induction and specification of the NC from the NPB are outlined. These networks clearly demonstrate the high degree of complexity of developmental GRNs. It becomes clear that (1) addition of further interactions based on experimental observations alone presents with certain limitations and is prone to inaccuracies and (2) conventional diagrammatic representations are no longer appropriate for the purposes of processing and evaluating these networks.

1.3.1. Neural plate border (NPB) induction

Prior to gastrulation, the neural and non-neural ectoderm become separated and poised to give rise to the central nervous system and the epidermis, respectively. Gradients of bone morphogenic protein (BMP) and fibroblast growth factor (FGF) signalling (Liem et al. 1995; Mayor, Guerrero, and Martínez 1997), as well as of retinoic acid (RA) and Wnt-mediated signalling (Kudoh, Wilson, and Dawid 2002; Takemoto et al. 2006) are responsible for this segregation (M. Simões-Costa et al. 2015; Groves and LaBonne 2014). These signals act as morphogens and emanate from different regions within the developing embryo. Posteriorly generated FGF signalling acts in concert with Wnt and BMP antagonists to specify neural ectoderm, a process which involves activation of pre-neural genes such as ERNI, Geminin, Otx2 and Sox2/3 (Mallamaci et al. 1996; Rex et al. 1997; Kroll et al. 1998; Andrea Streit et al. 2000). Conversely, Wnt and BMP signalling molecules, emanating from the lateral regions of the developing embryo, and antagonised by their respective medially-secreted antagonists, bias cells to adopt non-

neural fates (Wilson et al. 2001). Downstream of Wnt and BMP signals, members of the Msx, Dlx, GATA, Foxi and Tfap2 (also known as Ap2) families of TFs become activated in the non-neural ectoderm (Suzuki, Ueno, and Hemmati-Brivanlou 1997; Pera, Stein, and Kessel 1999; Sheng and Stern 1999; Hans et al. 2007; Hoffman et al. 2007).

On either side of the neural plate, the NPB separates it from the prospective epidermis. In order to establish the NPB and to gradually separate it into regions that will give rise either to NC or to PPE, it has been shown that spatiotemporally regulated levels of BMP, FGF and Wnt signalling play vital roles by inducing expression of NPB specifier genes (Fig. 1.7) (A Streit and Stern 1999; Steventon et al. 2009; Kwon et al. 2010; M. Simões-Costa et al. 2015; Schumacher et al. 2011; Yardley and García-Castro 2012; Garnett et al. 2012; Steventon and Mayor 2012). Expression of the genes Pax3, Tfap2a, Msx1, Zic1 and Hairy2 has been demonstrated in the NPB and the regulatory relationships between them, as well as their ability to regulate BMP, FGF and Wnt signalling in order to induce NC specifier genes while repressing adjacent fates have been extensively studied (Groves and LaBonne 2014; M. Simões-Costa et al. 2015).

In *Xenopus*, Wnt signalling has been shown to directly regulate the gene Gbx2, which cooperatively with Zic1 acts upstream of the NC-specific genes Pax3 and Msx1 (B. Li et al. 2009). Furthermore, Wnt signalling is reportedly upstream of the proto-oncogene c-Myc, which is expressed prior to other NC specifiers in the prospective NC domain (He et al. 1998; Bellmeyer et al. 2003). In the same system, Tfap2 was proven to be an essential mediator of FGF and Wnt signalling, a direct regulator of Pax3 and a direct or indirect activator of Msx1 and Hairy2 (de Crozé, Maczkowiak, and Monsoro-Burq 2011). Moreover, appropriately regulated levels of BMP signalling were found to induce expression of Zic1 (Tropepe et al. 2006), which is able to induce PPE by upregulating the marker gene Six1 (C.-S. Hong and Saint-Jeannet 2007; Marchal et al. 2009). This effect is reportedly inhibited through both Pax3 and Gbx2 (C.-S. Hong and Saint-Jeannet 2007; B. Li et al. 2009). In zebrafish, BMP signalling was shown to upregulate *prdm1a*, which is expressed in the NPB (Hernandez-Lagunas et al. 2005). Irreversible establishment of the NPB fate is thought to depend upon interactions between NPB specifiers (de Crozé, Maczkowiak, and Monsoro-Burq 2011). In support of this view, it has been indicated that Msx1 regulates Pax3 (Monsoro-Burq, Wang, and Harland 2005).

An interesting interplay is thus in place, according to which appropriate levels of Wnt, BMP and FGF induce the NPB by regulating the expression of Gbx2, Tfap2, Zic1, Pax3 and c-Myc, of which Gbx2, Tfap2a and Zic1 are widely expressed during gastrulation and Tfap2a as well as Zic1 function to specify non-NC fates (W. Li and Cornell 2007; B. Li et al. 2009). Intergenic interactions between initial NPB specifiers are vital for expression of Msx1 and Hairy2. Intriguingly, during early stages of NPB induction,

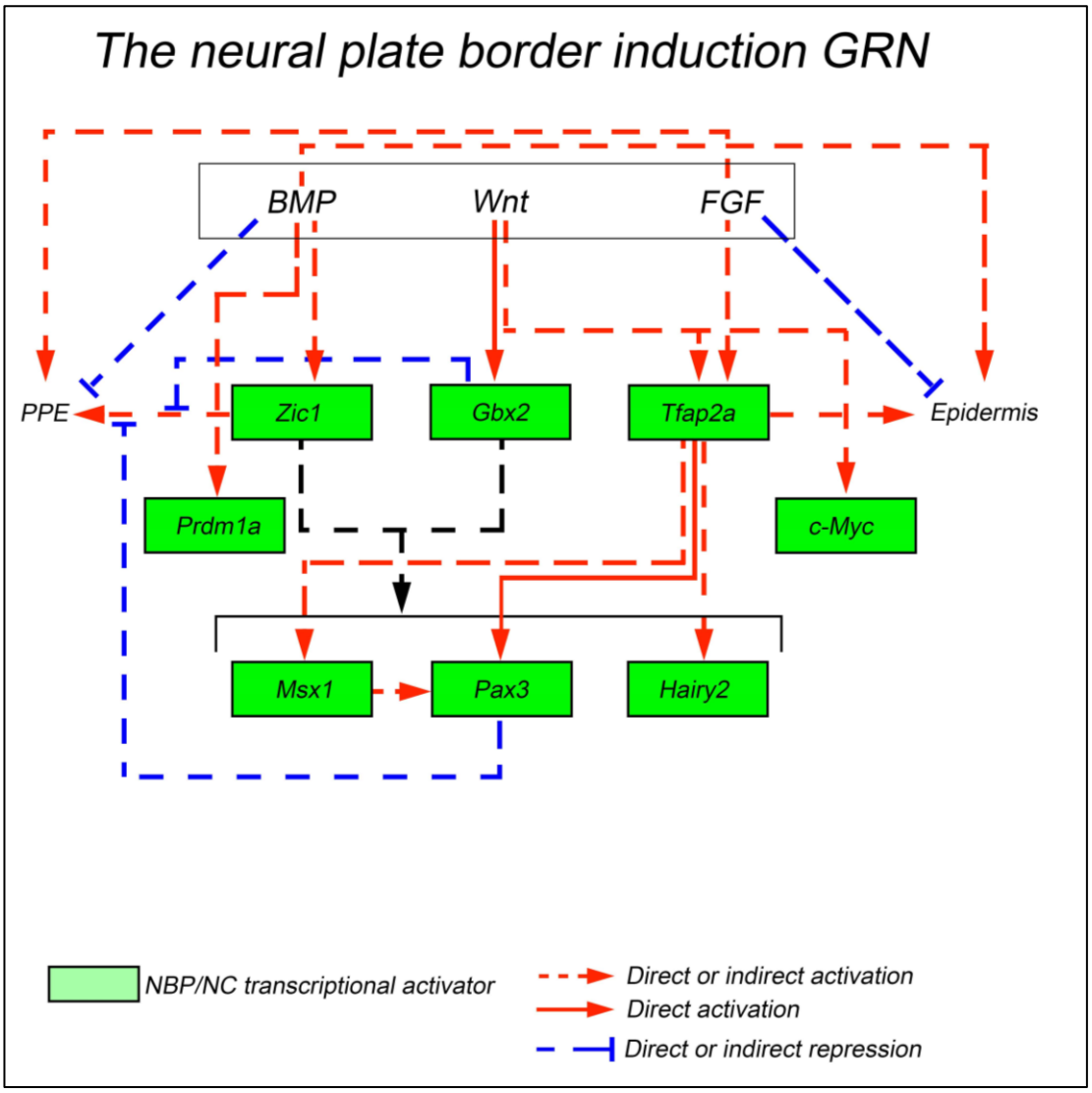


Figure 1.7. The NPB induction GRN. Diagrammatic description of how early morphogen signals (blank rectangle) emanating from different regions of vertebrate embryos regulate NPB-inducing and, subsequently, NC specifier genes (green rectangles), while repressing adjacent fates (epidermal and PPE). Although the order at which the genes become activated is largely unknown, the diagram aims to depict temporally antecedent interactions at the top, with downstream interactions progressively towards the bottom. The black arrow indicates cooperative activation between Gbx2 and Zic1 to induce all three NC-specific genes (Msx1, Pax3 and Hairy2). See main text for references.

expression of non- NC genes such as Dlx, GATA and Foxi family TFs has been indicated within the domain (Khudyakov and Bronner-Fraser 2009; M. Simões-Costa et al. 2015). Furthermore, loss of early NPB gene function leads to disruptions in both neural and non-neural ectoderm lineages. These observations highlight that at least during early stages, NPB induction does not correlate with irreversible fate segregation, but more likely with priming of the NPB to respond to lineage-specific signals (Groves and LaBonne 2014).

1.3.2. NC specification

The NPB specifier GRN functions to repress adjacent fates and to activate genes referred to as NCC specifiers, which proceed to definitively and irreversibly establish NCC progenitor fate by forming an intricate network of interactions (Fig. 1.8). Key players of the NC specification GRN are the TFs Foxd3, Snai1/2, Sox9/10, c-Myc, Myb, Ets1 and Twist, while the NPB specifiers Tfap2a, Pax3, Zic1 and Msx1 continue playing important roles (Khudyakov and Bronner-Fraser 2009; Groves and LaBonne 2014; M. Simões-Costa et al. 2015). It is important to note that Wnt function remains crucial during NC specification, as it allows Pax3 and Zic1 to form a positive feedback loop (Sato, Sasai, and Sasai 2005).

One key feature of the derived GRN is the presence of positive feedback loops and of a high degree of functional redundancy. These properties, as well as self-regulation of TFs such as Ets1 and Pax3, ensure NCC fate maintenance even if the function of one or two players is eliminated (Sauka-Spengler and Bronner-Fraser 2008). Moreover, maintenance of multipotency is an important feature of early NCCs. In several species, Foxd3 has been shown to maintain NC fate by repressing differentiation genes, for instance Mitf, and is downregulated in specified and migrating NC derivatives (Kos et al. 2001; Dottori et al. 2001; Ignatius et al. 2008; Teng et al. 2008). Finally, in chick embryos it was demonstrated that cNCC versus tNCC GRNs slightly differ. For instance, the effects of Ets1 appear to be restricted to NC specification in cNCCs, with its functions being accomplished by Zic1 and other factors in tNCCs (M. S. Simões-Costa et al. 2012; Barembaum and Bronner 2013).

Initiation of the NC specification GRN depends upon the balance between Pax3 and Zic1, which is established in a Wnt, BMP and FGF dependent manner, as described in the previous section (C.-S. Hong and Saint-Jeannet 2007). Pax3 and Zic1 form a positive feedback loop (Sato, Sasai, and Sasai 2005; Plouhinec et al. 2014), which is reinforced by Msx1-dependent upregulation of Pax3 (Monsoro-Burq, Wang, and Harland 2005). The definitive NC marker Foxd3 (Sasai, Mizuseki, and Sasai 2001; Stewart et al. 2006) has been shown to become upregulated via a complex mechanism in the chick NC. Both

Pax7 and Msx1/2 were shown to directly bind regulatory elements on the Foxd3 promoter, although in both cases upregulation of the gene required additional direct binding of Ets1 in the cNC or Zic1 in the tNC (M. S. Simões-Costa et al. 2012). Upregulation of Foxd3 by Snai1 (Aybar, Nieto, and Mayor 2003) as well as in a Wnt-dependent manner as a result of the Pax3/Zic1 feedback loop (Sato, Sasai, and Sasai 2005) was further demonstrated in *Xenopus*. In zebrafish, the NPB specifier *pdrm1a* was shown to directly activate the expression of both *foxd3* and *tfap2a* (Powell et al. 2013). Furthermore, both c-Myc and Pax3 play important roles in NC specification in *Xenopus* by mediating Tfap2a expression, which is indispensible for Pax3 and Zic1 dependent upregulation of NC specifiers such as Foxd3 and Snai2 (de Crozé, Maczkowiak, and Monsoro-Burq 2011).

Direct upregulation of Snai1 through both Pax3 and Zic1 was demonstrated in *Xenopus* (Sato, Sasai, and Sasai 2005; Plouhinec et al. 2014). Pax7 was also shown to either indirectly or directly upregulate Ets1 expression in the cNC (Barembaum and Bronner 2013). Additionally, likely indirect activation of Ets1, Snai2, Twist1 and Sox10 by the transcriptional repressor Snai1 was suggested (Aybar, Nieto, and Mayor 2003; Honoré, Aybar, and Mayor 2003). In zebrafish, *foxd3* and *tfap2a* positively regulate *sox9* and *sox10* (Arduini, Bosse, and Henion 2009). This effect only became clear upon loss of *foxd3* and *tfap2a* function simultaneously, which is likely the result of redundancy and positive feedback within the GRN overriding loss of either *tfap2a* or of *foxd3* individually. Moreover, Sox10-mediated Snai2 activation was shown in *Xenopus* (Honoré, Aybar, and Mayor 2003), while Sox10 itself was found to be directly regulated by cMyb in the cNC of chick embryos (P. Betancur, Bronner-Fraser, and Sauka-Spengler 2010). Notably, studies in chick have shown that Sox9 is able to activate expression of both Foxd3 and Sox10 (Cheung and Briscoe 2003; P. Betancur, Bronner-Fraser, and Sauka-Spengler 2010), while it directly binds the Ets1 promoter and upregulates Ets1 expression in cNCCs (Barembaum and Bronner 2013). Upregulation of Sox10 by Sox9 is direct in cNC but might occur either directly or indirectly in the tNC.

Finally, the NC inducer and specifier gene Hairy2 was demonstrated to maintain the mitotic state of NCCs by repressing the cyclin-dependent kinase inhibitor $p27^{xic1}$ in *Xenopus* embryos (Nagatomo and Hashimoto 2007). The role of Hairy2 in NC development is controversial, with a body of data suggesting that Bmp4 expression is repressed by Hairy2 (Glavic et al. 2004) and that this interaction mediates downregulation of Id3, a factor which allows cells to proceed to fate restriction and differentiation stages (Nichane et al. 2008). However, morpholino experiments conducted using the same model organism contradict the role of Hairy2 in Bmp4 repression (Nagatomo and Hashimoto 2007).

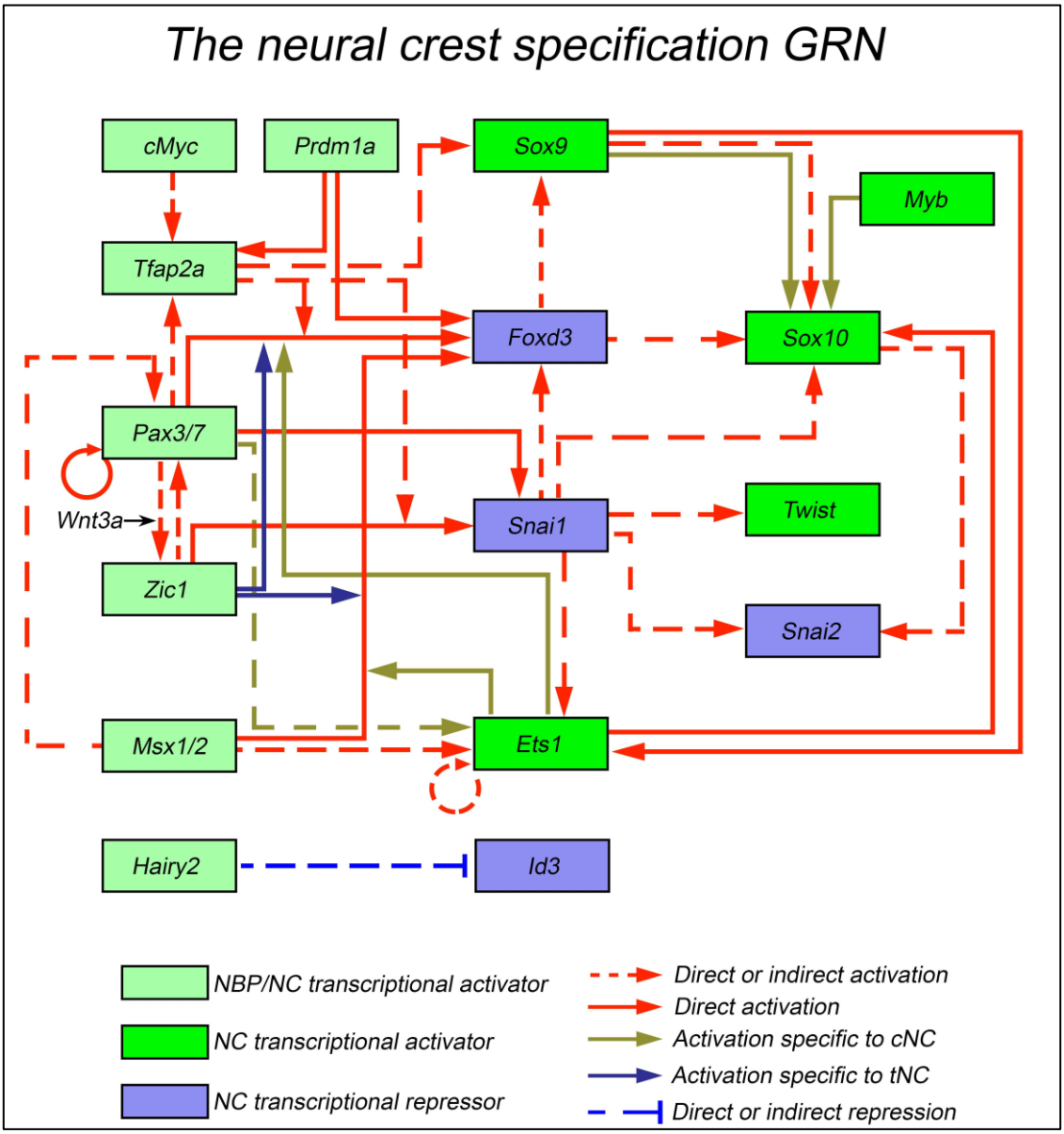


Figure 1.8. The NC specification GRN. The diagram illustrates interactions taking place in specified NCCs. Due to their likely redundant functions across species, Pax3 and Pax7 as well as Msx1 and Msx2 are not differentiated. NPB inducers (light green) continue playing important roles in the specification GRN. Transcriptional repressors are often reported to upregulate downstream factors, likely through suppressing unidentified repressors of those factors. Where an arrowhead is placed along the length of another arrow, presence of two different upstream regulators is necessary for the corresponding transcriptional effect to take place. See main text for references.

The above correspond to major regulatory interactions indicated by thus far published experimental data. The presented NC specification GRN functions not only to achieve specification of NCCs and maintenance of multipotency, but also towards activation of the transcriptional programs that will allow NCCs to delaminate from the dorsal surface of the closed neural tube prior to entering the migratory pathways. This process requires cytoskeletal changes (Dent, Polson, and Klymkowsky 1989) and fine regulation of intercellular interactions (Taneyhill and Schiffmacher 2013). Furthermore, gene regulatory interactions within NCCs render them able to invade tissues in response to chemotactic signals and to undergo sequential fate restriction processes before giving rise to differentiated derivatives. GRNs governing these processes will not be discussed in the present review.

1.3.3. Developmental GRNs: advances and limitations

To construct GRNs such as the ones presented, traditional experimental approaches have primarily been employed. Loss of function studies, in which morpholino technology was employed to knock-down genes of interest, represent the primary sources of evidence for intergenic interactions. Due to technical constraints commonly associated with generation of mutant lines, a significant number of these studies were carried out taking advantage of the morpholino technology, resulting in knock-down of a gene of interest. However, morpholino-mediated loss of gene function has been implicated in non-specific targeting and in induction of the p53 pathway which activates precocious apoptosis (Robu et al. 2007; Eisen and Smith 2008; Schulte-Merker and Stainier 2014; Rossi et al. 2015). Furthermore, loss of function studies are unable to indicate direct versus indirect interactions within the GRNs. Combined, this may lead to the addition of redundant or otherwise inaccurate links to the network, as a result of not fully understanding the molecular interactions underlying a phenotype.

A large body of work has been built on conducting overexpression studies, in which a certain gene or combination of genes is ectopically expressed and the downstream consequences on the expression of potential target genes are assessed. Although such an approach has proven very useful to derive primarily indirect but in certain cases direct intergenic interactions (Plouhinec et al. 2014), results must be interpreted cautiously and validated in order to avoid incorporating non-specific effects in GRNs. Non-specific effects commonly arise due to ectopically introducing non-physiological levels of transcripts. Such an intervention may result in abnormal activation of regulatory elements either responding to a related TF or, particularly in the case of enhancer elements, present with different degrees of accessibility in different contexts.

In recent years, transcriptomic studies have significantly expanded the lists of genes to be studied and added to existing GRNs (Simoes-Costa et al. 2014). Such studies rely upon generation of novel transgenic lines that express fluorescent proteins in specified cell populations of interest, allowing for isolating this population via fluorescence activated cell sorting (FACS). Moreover, identification of active regulatory elements, for example the so-called super-enhancers, in cell populations (Kaufman et al. 2016) and the increasingly widespread usage of chromatin immunoprecipitation and sequencing (ChIP-seq) studies (P. Betancur, Bronner-Fraser, and Sauka-Spengler 2010; M. S. Simões-Costa et al. 2012; Furey 2012; Gorkin et al. 2012) have proven invaluable for elucidating the molecular mechanisms guiding developmental processes and for identifying likely direct interactions. In summary, both traditional, well-established genetic techniques, as well as newly developed tools aid elucidation of molecular mechanisms and construction of developmental GRNs. Despite the sophisticated toolbox available, further complications present when analysing and interpreting the resulting GRNs.

Experimentally derived GRNs, such as those regulating NC development, can prove difficult to interpret using diagrammatic representations, due to their complexity (Fig. 1.7; Fig. 1.8) (Baxter et al. 2010; M. Simões-Costa et al. 2015). Furthermore, it becomes apparent that different studies identify interactions which might be inappropriately incorporated within GRNs. An immediate example is the depicted upregulation of Snai2 independently by Snai1 and by Sox10, on the basis of experimentally determined loss of Snai2 expression in relevant loss of function studies, which also suggested that Sox10 was itself activated by Snai1 (Fig. 1.8) (Aybar, Nieto, and Mayor 2003; Honoré, Aybar, and Mayor 2003). When diagramatically presenting this set of interactions, it would be expected that Snai1 is capable of maintaining Snai2 expression even in the absence of Sox10. This, however, was experimentally shown to not be the case, as knocking-down Sox10 in *Xenopus* embryos results in robust downregulation of Snai2 (Honoré, Aybar, and Mayor 2003). In this example of a simple and relatively isolated set of interactions, the arising inconsistency between gene expression dynamics, predicted based on the GRN's structure, and experimental observations can be easily discerned. Nevertheless, less easily identifiable misinterpretations of experimental data have likely occurred when generating developmental GRNs. The need for a more robust, predictive method of analysing and evaluating networks thus emerges. Systems biology approaches, in particular mathematical modelling of GRNs, is increasingly appreciated as it allows for accurately describing data and deriving predictions as well as experimentally testable hypotheses (Gunawardena 2014). Examples of GRNs which benefitted from mathematical modelling are described in the following section.

1.4. Using systems biology to construct GRNs

1.4.1. Mathematically modelling GRNs

Mathematical modelling in developmental biology was first introduced by Alan Turing, who developed models able to simulate morphogenesis using the so-called reaction-diffusion differential equations (Turing 1952). Mathematical modelling allows for rigorously testing existing hypotheses as well as for deriving testable ones that can then be evaluated using experimentation (Tomlin and Axelrod 2007; Gunawardena 2014). Initially, mathematical models were employed solely for the purpose of testing assumptions, since experimental data was limited and little knowledge was available in regards to molecular mechanisms underlying interactions. More recently, whole genome sequencing and sophisticated approaches for identifying candidate components of a system, such as those mentioned in section 1.3.3 for developmental GRNs, aided the construction of more detailed and mechanistically accurate predictive networks (Tomlin and Axelrod 2007). In addition to advances in experimentation methodologies, increased computational processing capacity now allows for integrating mathematical models describing small-scale systems to produce predictions with regards to complex systems, for example processes and responses occurring at the tissue level as opposed to the cellular level (Tomlin and Axelrod 2007).

There are two major types of mathematical models that are employed to describe biological process: discrete and continuous models (Wittmann et al. 2009; Ballerstein et al. 2013). Discrete and continuous models each offer distinct advantages and disadvantages when mathematically describing biological processes. Discrete models are strictly qualitative and do not traditionally encompass the element of time (Ballerstein et al. 2013). As suggested by the name, these models incorporate variables that can only adopt discrete states. In the case of GRNs these states would correspond to genes being either expressed or not expressed. The major advantage of discrete models is the lack of requirement for detailed knowledge on concentrations and parameter values, such as affinity constants. These models are also simplified by the exclusion of the element of time, which is however considered a disadvantage when attempting to capture the continuous dynamics of developmental processes in particular.

Continuous models employ systems of differential equations to simulate the rate of change of a variable over time. Differential equations are functions that relate a quantity to its derivative, a term representing the rate of change of that quantity. The terms ordinary and partial differential equations (ODEs or PDEs) relate to the number of independent variables in each case (Agarwal and O'Regan 2009). Only one independent variable is present in ODEs, whereas PDEs incorporate more than one variables. At a more advanced level, systems of differential equations describe rates of change of

different components over time, when these components influence each other. Important limitations associated with the use of ODEs concern the choice of constant parameters, which are often hard to derive experimentally. In this study and in many others from the published literature (Greenhill et al. 2011; Balaskas et al. 2012; Cohen et al. 2014), systems of ODEs have been used to mathematically model developmental GRNs. Limitations in regards to parameter choice are addressed differently in different studies, for example by conducting sensitivity tests (Balaskas et al. 2012).

A major limitation when using deterministic mathematical models to describe a network relates to the assumption that the system in question is neither influenced by random variations of its own components, nor by external factors. In other words, deterministic models assume that a specific input will always produce the same corresponding output according to the used set of equations. To take into account the intrinsic variations and randomness inherently present in biological and chemical systems, stochastic mathematical models accompanied by the corresponding advanced computational methods, such as the Gillespie algorithm, have been developed (Gillespie 1977; Ditlevsen and Samson 2013). In this work, deterministic systems of ODEs were exclusively employed to describe the GRN underlying iridophore development, therefore stochastic mathematical models will not be discussed further in this section. Nevertheless, it is important to consider applying stochastic mathematical modelling during subsequent steps of GRN construction to identify potential effects of randomness in both gene expression and TF interaction on fate decision of multipotent progenitors, as well as on the generation of escaper cells in mutant contexts (refer to sections 1.5.1 and 1.5.3).

1.4.2. Examples of mathematically modelling GRNs

Both discrete and continuous mathematical models have been productively used to simulate dynamics in developmental GRNs. Three examples of such endeavours are briefly outlined in this section. First, modelling of GRNs functioning in HSCs using discrete modelling. Second, investigating GRNs initiated by Sonic hedgehog signalling, which function to pattern the spinal cord. Third, generating the GRN governing melanophore development. Systems of ODEs are used both in the second and third example.

1.4.2.1. Haematopoietic Stem Cells (HSCs)

HSCs are responsible for generating the entire complement of blood cells throughout the life of an individual, starting from embryogenesis (Orkin and Zon 2008). Prior to giving rise to differentiated cell types, including red blood cells, platelets, granulocytes, macrophages and lymphocytes, HSCs undergo a series of partial fate restriction

processes to generate common precursors of the myeloid and the lymphoid lineages. The common myeloid progenitor has further been reported to generate a bipotent megakaryocyte/erythroid progenitor and a tetrapotent granulocyte/macrophage progenitor (Orkin and Zon 2008). Although several studies have focused on mathematically modelling the GRN governing HSC differentiation towards their different derivatives, here two representative ones are highlighted to illustrate the predictive power of systems biology.

In the first study, Krumsiek and colleagues used Boolean regulatory logic to construct the GRN governing the dynamics of a particular stage of haematopoiesis, namely the differentiation of HSC-derived common myeloid progenitors towards four distinct derivatives, erythrocytes, megakaryocytes, monocytes and granulocytes (Krumsiek et al. 2011). Boolean networks fall under the category of discrete mathematical models. Since only qualitative understanding of genetic interactions was available (for example the knowledge that gene A upregulates gene B), Boolean equations were used to describe each relationship (Krumsiek et al. 2011). The constructed GRN accurately predicted the segregation of four differentiated lineages, each displaying the experimentally established genetic signature, from common myeloid progenitors. Importantly, even when the network was perturbed by inducing loss of gene function *in silico*, the four cell lineages were predicted to display phenotypes consistent with those observed experimentally. This type of mathematical modelling yielded a number of biologically testable hypotheses, the validity of which was not, however, subsequently tested. Although in this context the Boolean approach appeared to accurately reproduce biological observations, a major limitation lies in its inability to simulate intermediate expression levels and time-dependent effects. As a result, fate restricted precursors derived from the common myeloid progenitor could not be identified in terms of their gene expression signatures.

Another study attempted discrete mathematical modelling of a regulatory module underlying HSC specification either towards the erythroid or the myeloid lineage (Tian and Smith-Miles 2014). Notably, it was demonstrated that addition of a novel gene in the network (GATA-2) produced biologically accurate predictions. Importantly, in this work the effects of stochasticity were investigated and a thorough investigation of parameter space was conducted using an appropriate algorithm.

1.4.2.2. Patterning of the spinal cord

During vertebrate embryogenesis, patterning of the spinal cord to aid generation of different types of neurons is mediated by a gradient of Sonic hedgehog (Shh) signalling, emanating from the ventral side of the embryo and regulating gene expression through

Gli effector proteins (Ribes and Briscoe 2009). Systems of ordinary differential equations have proven instrumental in describing the mechanism by which the Shh morphogen gradient is able to determine lineage specification at different levels of the neural tube.

This type of continuous mathematical modelling played an important role in revealing various overlooked mechanisms. For instance, the model predicted that a gradient of Shh is not sufficient to induce the observed expression patterns in the neural tube. Instead, it was proposed and experimentally validated that the established connectivity between members of the Shh-triggered GRN is essential for patterning to occur (Balaskas et al. 2012). Furthermore, intriguing predictions concerned the hysteresis effect, referring to the acquisition of prolonged cellular memory regarding the level and duration of the received signal, which was subsequently validated experimentally (Balaskas et al. 2012). Moreover, elegant studies employed mathematical modelling and Bayesian computational methods to address the molecular basis of the experimentally observed adaptation response, in which cells continuously exposed to Shh signalling gradually cease to respond (Cohen et al. 2015).

Overall, in the context of spinal cord development, mathematical modelling aided in depth theoretical investigation of the involved dynamics in time and space. Theoretical testing, followed by experimental validation of proposed hypotheses shed light on the mechanisms of spinal cord patterning, which was proven to require mechanisms more complex than the widely accepted effects of morphogen concentration gradients.

1.4.2.3. Melanophore development

In zebrafish, systems of ODEs were effectively used to elucidate the GRN governing melanophore development (Greenhill et al. 2011). In this study, experimental techniques were used to derive a simple, preliminary GRN (model A), which was mathematically modelled and computationally simulated to predict gene expression dynamics over time, for each gene of the network. Predictions from the model were experimentally evaluated and testable hypotheses were formulated in order to render predictions more biologically accurate. These hypotheses, which were first theoretically tested using the mathematical model, involved either the addition of novel factors to the network or alterations to the connectivity. Promising hypotheses were then experimentally tested prior to incorporation to the existing network. Thus, iterations of theoretical testing of novel hypotheses and experimental validation led to refining and expanding model A to ultimately generate a more sophisticated model C, demonstrated to more accurately reflect biologically determined gene expression dynamics.

Although this methodology was successful in evolving the zebrafish melanophore GRN, certain limitations were recognised. Firstly, due to lack of quantitative gene expression

data in the developing melanophore lineage, predictions could only be evaluated qualitatively. Moreover, the parameters chosen to implement the system of ODEs were arbitrarily chosen. This problem was tackled by performing computational explorations of parameter space, in which, however, only two parameters could be varied at a time (Greenhill et al. 2011).

This iterative process of mathematical modelling and experimentation provides a novel way of processing complex GRNs. Networks such as those governing NC development (refer to section 1.3) could potentially benefit from this approach, as they have become too complex to analyse, evaluate and evolve using experimental methods alone.

1.4.2.4. Caveats of mathematical modelling

When considering mathematical modelling as a tool to improve our understanding of a complex system, the caveats of such an approach should be noted. Specifically, obtaining reliable predictions requires mathematical representation, using discrete, continuous or stochastic mathematical models, as discussed above, of well-established genetic interactions, derived from reliable experimental data or reasonable hypotheses (Gunawardena 2014). In certain cases, key features of a network could be overlooked, for example due to lack of appropriate technical resources. Furthermore, aspects of a system under investigation are almost always disregarded, either due to lack of knowledge or on the basis of being beyond the scope of the network in question. As becomes clear in the case of GRNs, mathematical modelling is only feasible by isolating and focusing on a process of interest. Therefore, one could conclude that all GRNs are inherently incorrect and, hence, could not be considered accurate representations of biological processes.

Although this concern is valid, the aforementioned examples demonstrate that mathematical modelling of biological systems can provide valuable insight when (1) the appropriate questions are asked, (2) the right abstractions are made to maintain only the necessary level of detail, and (3) the model is falsifiable, in other words when assumptions are exposed and become subject to being refuted (Gunawardena 2014). This can be achieved by methodically evaluating predictions of the model via experimentation.

Furthermore, a concern when building mathematical models aiming to describe molecular processes, is with regard to how the unavoidable uncertainty is approached (Kirk, Babtie, and Stumpf 2015). In particular, the inability to quantitatively determine the abundances of all factors within a system and to measure the kinetics of their interactions leads to considerable uncertainty, not only when evaluating predictions and deriving parameter constants but, importantly, when determining the structure of the model.

Specifically, it may be concluded that a given network accurately reflects biology, on the basis of its predictions fitting the data, while alternative networks, which are being disregarded, could also produce appropriate predictions. It has been noted that this lack of confidence in many crucial aspects of the investigated models often remains unacknowledged, leading to over-interpretation of conclusions (Kirk, Babtie, and Stumpf 2015). Methods to more comprehensively address uncertainty in mathematical models are being developed. For instance, network inference is used to reverse-engineer systems and construct possible networks that fit experimentally determined datasets (Crombach et al. 2012; Villaverde and Banga 2013).

1.5. The GRN governing iridophore development

In contrast to the melanophore GRN, the one underlying iridophore specification from NCC progenitors has not been identified. As a consequence, our understanding of NC development and fate segregation is incomplete, especially when considering the melano-iridoblast. To begin to establish the iridophore GRN, candidate genes of interest were identified from the existing literature, prior to determining their interactions using mathematical modelling and experimentation.

To identify genes important for iridophore development, mutagenesis screens (R N Kelsh et al. 1996), small or large-scale gene expression studies (Thisse et al. 2001; James A Lister et al. 2011) and large-scale screening for factors associated with iridophore development (Higdon, Mitra, and Johnson 2013) were considered. Candidate genes were selected according to the following criteria: (1) their gene expression patterns in relation to the positions of iridoblasts and iridophores and (2) their mutant phenotypes. Specifically, mutants of genes of interest were required to present with iridophore phenotypes, namely either reduction or increase of iridophore numbers.

1.5.1. SoxE factors

There are three SoxE TFs in vertebrates: Sox8, Sox9 and Sox10. They contain a DNA-binding HMG domain, a dimerization and a transactivation domain, as well as a central protein interaction domain (K2) (Weider and Wegner 2016). The important roles of SoxE factors during the processes of NC specification and of development of PNS derivatives has been widely investigated (Peirano and Wegner 2000; Cheung et al. 2005; C. Hong 2005; Carney et al. 2006). SoxE proteins interact in a cell and stage specific manner with other TFs, such as Pax3 and Olig2, to cooperatively expression of downstream targets (Bondurand et al. 2000; Weider and Wegner 2016). Furthermore, SoxE factors have been shown to recruit chromatin modifiers to regulatory elements of their targets (Weider and Wegner 2016).

In regards to iridophore development, expression of *sox8* has not been demonstrated in iridophore locations (Yan et al. 2005), thus the gene is an unlikely candidate for the GRN. Although expression of the *Sox9* orthologue, *sox9b,* has not been indicated either in pigment cell precursors or within mature pigment cells (Thisse et al. 2001; Yan et al. 2005), loss of *sox9b* function resulted in elimination of iridophores accompanied by enlargement of melanophores at 4 dpf (Yan et al. 2005). Even though this was an intriguing result, the mutant allele used to generate these results (*sox9b*^{b971}) corresponds to a terminal deletion removing 10 cM from the LG3 chromosome. Because a large number of genes was deleted, which could have influenced the resulting phenotype, and due to lack of *sox9b* expression in positions relevant for the iridophore lineage (Yan et al. 2005; Chuang et al. 2010), the gene was excluded from the GRN.

Genome-wide mutagenesis screening identified the TF *sox10* as a key player for iridophore development (R N Kelsh et al. 1996). Similarly to mouse models, loss of *sox10* function in zebrafish results in embryonic lethality preceded by megacolon agangliosis, lack of pigment cells and defects in PNS derivatives. However, no effects have been observed in the craniofacial cartilages (R N Kelsh and Eisen 2000; R.N. Kelsh 2006). Upon loss of *sox10* function, chromatophore derivatives of the tNC fail to migrate, while it was indicated that neurons and glia contributing to the DRGs enter the migratory routes normally (R N Kelsh, Schmid, and Eisen 2000; Dutton et al. 2001). In humans, loss of SOX10 function results in a variety of disorders, including Waardenburg-Shah syndrome and Hirschsprung disease (Bondurand et al. 2000; R.N. Kelsh 2006).

Zebrafish *sox10* mutant embryos (also referred to as *colourless*) display lack of all three chromatophores. In some mutant embryos, very few differentiated cells are observable. These are referred to as ‘escaper’ cells and are believed to arise as a result of other transcriptional regulators functioning redundantly with *sox10*. As the number of escaper cells is typically very low, it is plausible that activation of the chromatophore GRNs by alternative regulators occurs in a stochastic manner, under specific sets of circumstances. Described *sox10* mutant alleles include *sox10*^{m618}, which harbours a point mutation leading to a non-conservative change within the HMG domain (Dutton et al. 2001) and *sox10*^{t3}, which arose from spontaneous insertion of a 1.4 kb transposon within the *sox10* open reading frame, resulting in a premature stop codon (R N Kelsh et al. 1996; Dutton et al. 2001). The *sox10*^{t3} mutation produces a more severe phenotype compared to *sox10*^{m618} and was thus used for loss of *sox10* function studies presented in later sections.

As described in section 1.1.4, expression studies performed on *sox10*^{t3} mutants indicated the existence of a NC-derived, partially fate-restricted pigment cell progenitor, the chromatoblast (Lopes et al. 2008). Specifically, in *sox10* mutants the iridophore gene,

ltk (refer to section 1.5.3), was prominently expressed in persisting cells of the premigratory NC domain, which however did not express markers of specified NCCs, such as *foxd3* and *snai2*. Further data suggested the existence of a tetrapotent progenitor, the chromatoglioblast (Nikaido, M., unpublished data). Work presented here builds on the concept of the chromatoglioblast, which undergoes progressive fate restriction processes to ultimately generate mature iridophores (Fig. 1.4).

Considering the established role of *sox10* in specification of NC derivatives through upregulation of master regulators such as the melanophore gene *mitfa* (Elworthy et al. 2003; R.N. Kelsh 2006), and the lack both of *ltk* positive migratory and mature iridophores (Lopes et al. 2008), sox10 was regarded as a vital player within the iridophore GRN.

1.5.2. Micropthalmia TF family (MiT)

Mitf, Tfe3, Tfeb and Tfec are members of the MiT family of highly conserved basic helix loop helix-leucine zipper (bHLH-LZ) TFs (James A Lister et al. 2011; Ploper and De Robertis 2015). These TFs have been shown to homo- and hetero-dimerise with each other, but not with other related bHLH transcriptional regulators such as c-Myc, Max and USF (Hemesath et al. 1994; Pogenberg et al. 2012). Interestingly, tissue-specific regulation and distinct functions of MiT family genes depend upon alternative splicing, utilisation of distinct 5' exons and production of truncated transcripts (Kuiper et al. 2004). Furthermore, it was reported that different tissues express MiT factors in particular ratios, which may affect heterodimerisation tendencies and, thus, regulation of downstream targets (Kuiper et al. 2004).

Across species, *Mitf* is the melanocyte master regulator and plays an important role in proliferation in the context of melanomas (Ploper and De Robertis 2015). In zebrafish, the two paralogues *mitfa* and *mitfb* function redundantly to promote melanogenesis (J A Lister, Close, and Raible 2001). *MITF* mutations in humans are known to cause Waardenburg syndrome type II and Tietz syndrome and have been associated with hypopigmentation of the skin, micropthalmia and deafness (Read and Newton 1997; Smith et al. 2000; Izumi et al. 2008). In zebrafish, *mitfa* is expressed in melanophore precursors from very early stages of specification (approximately 21 hpf). *sox10* directly upregulates *mitfa* by directly binding to its promoter in zebrafish, but also in mammals (Potterf et al. 2000; Elworthy et al. 2003). *mitfa*^{w2} mutants (also referred to as *nacre* or *nac*^{w2}), which were used in this study harbour a nonsense mutation and present with lack of embryonic melanophores (J A Lister et al. 1999), with the exception of those of the eye, where *mitfb* and *mitfa* appear to function redundantly (J A Lister, Close, and Raible 2001).

Intriguingly for the purposes of the iridophore GRN, *mitfa*^{w2} mutants display a 40% increase in mature iridophores (J A Lister et al. 1999). Moreover, it has been shown that repression of *mitfa* by *foxd3* (discussed in section 1.5.4) is responsible for driving iridophore instead of melanophore fate from a common progenitor (Curran, Raible, and Lister 2009; Curran et al. 2010). Due to its proposed role in the melanoiridoblast and the effects observed in *mitfa*^{w2} in the iridophore lineage, the gene was considered another player of the GRN. Importantly, its rapid downregulation in the iridophore lineage soon after iridophore specification is expected due to its key role as the melanophore master regulator.

The zebrafish genome contains two *tfe3* paralogues, *tfe3a* and *tfe3b* (James A Lister et al. 2011). Based on their established expression patterns, none of *tfe3a*, *tfe3b* or *tfeb* are expressed in zebrafish during NC and pigment cell development (James A Lister et al. 2011). However, *tfec* expression has been demonstrated both in premigratory NCCs, and in mature iridophores at different stages of development (James A Lister et al. 2011). In addition, morpholino-mediated loss of *tfec* function has been shown to have a severe effect on the iridophore lineage, with iridophores becoming completely abolished in *tfec* morphants (J. Lister unpublished data). Taking into account the high degree of evolutionary similarity between major domains of the master regulator *mitfa* and *tfec*, and considering the expression pattern and loss of function observations for each gene with regards to the melanophore and the iridophore lineage, respectively, *tfec* is here considered an excellent candidate for the master regulator of the iridophore lineage. To establish this role of *tfec*, further studies are required to demonstrate that its presence is not only necessary but also sufficient for iridophore development.

1.5.3. Transmembrane receptors

Ltk is an insulin-like RTK expressed in mammalian B-cells and neuronal tissues (Ben-Neriah and Bauskin 1988; Bernards and de la Monte 1990; J. B. Weiss et al. 2012). Gain of function SNPs in the human *LTK* gene were shown to confer susceptibility to the complex disease systemic lupus erythematosus (N. Li et al. 2003). Moreover, LTK expression has been reported in human leukemias (Maru, Hirai, and Takaku 1990), however the exact functions and downstream targets of ltk remain unidentified. Furthermore, Ltk has remained an orphan receptor for an extensive period of time (Lemke 2015) and it is only recently that efforts from several groups have focused on identifying the endogenous ligands of Ltk. The favoured and widely supported candidates are FAM150A and FAM150B in zebrafish (FAM150 in mammals) (Zhang et al. 2014; Guan et al. 2015; Reshetnyak et al. 2015).

Several zebrafish mutants presenting with either loss or gain of *ltk* function have been generated. *ltk*^{ty9} and *ltk*^{ty82} are amongst commonly used loss of function mutants (also referred to as *shady* or *shd*), presenting with reduced or absent embryonic and adult iridophores, but not with defects in the development of other chromatophore lineages (Lopes et al. 2008). *ltk*^{ty9} homozygous mutants display partial loss of iridophores and can survive to adulthood. *ltk*^{ty82} mutants, solely utilised in this study, bear a nonsense mutation leading to truncation of the receptor prior to the tyrosine kinase domain. Embryos present with severe lack of iridophores and do not survive into adulthood. Specifically, homozygous *ltk*^{ty82} mutant embryos maintain a maximum of 3 escaper iridophores along the trunk and tail (Lopes et al. 2008). With regards to gain of function mutants, *ltk*^{mne} (*moonstone* mutant line), bears a missense mutation resulting in a dominant hyperactive form of *ltk* (Fadeev et al. 2016). Both embryonic and adult iridophores of *ltk*^{mne} mutants are increased, but other pigment cells do not appear affected, except for melanophore numbers on adult scales. Using this mutant line the role of *ltk* in several aspects of *adult* iridophore development, including proliferation, survival and homotypic iridophore interactions were indicated (Fadeev et al. 2016).

Another class of receptors involved in iridophore development are endothelin receptors (Ednr), which are seven pass G-protein coupled receptors stimulated by extracellular endothelin (Edn) ligands. Mice bearing mutations that affect endothelin signalling mediated by the Ednrb receptor paralogue present with aganglionic megacolon and pigmentation defects, typical of Hirschsprung disease (Baynash et al. 1994; Hosoda et al. 1994). In humans, a subset of Hirschsprung disease cases as well as hyperpigmentation conditions and cancer result from mutations in EDNRB-dependent signalling (McCallion and Chakravarti 2001; Kadono et al. 2001; H. J. Lee, Wall, and Chen 2008). The importance of endothelin signalling in melanocyte development and melanoma formation has thus been established and carefully investigated (Saldana-Caboverde and Kos 2010).

Studies in zebrafish have identified the mammalian Ednrb orthologue, *ednrb1*, as an essential gene for development and patterning of all three pigment cell lineages at late stages of metamorphosis (Parichy et al. 2010). Intriguingly, although *ednrb1* is expressed in pigment cell precursors during embryonic development, embryos presenting with loss of receptor function (referred to as *rose* mutants) present with no defects during early pigment cell development. Possible explanations for this effect during embryogenesis could be functional redundancy, compensation, or effects on protein stability, for instance resulting from post-transcriptional modifications, which would render it non-functional (Parichy et al. 2010). However, mutant embryos fail to develop normal chromatophore numbers following metamorphosis, and to acquire the

WT adult pigment pattern, suggesting an important role of *ednrb1* in adult pigment lineages, including iridophores. Recently, this notion was further investigated and the Ednrb1 ligand was described (Krauss et al. 2014). In this study, *ednrb1* was not included in the iridophore GRN due to the apparent lack of functionality during development of embryonic iridophores.

1.5.4. Transcriptional repressors

Inhibitor of DNA binding (or inhibitor of differentiation) 2a (*id2a*) is another negative regulator of transcription which has been implicated in iridophore development (F. Rodrigues, PhD thesis). Id family TFs are HLH proteins, able to form homodimers and to heterodimerise with bHLH factors. bHLH molecules, which are capable of DNA-binding through their basic domain, thus become sequestered and their effects on transcription are inhibited (Ling, Kang, and Sun 2014). Id factors, four of which have been identified in mammals (Id1-Id4), have been suggested to promote stem cell proliferation and multipotency, while repressing differentiation towards mature lineages in several systems, particularly in the mammalian CNS (Ling, Kang, and Sun 2014; Diotel et al. 2015).

In zebrafish, few studies address the expression, function and genetic regulation of id proteins (Dickmeis et al. 2002; Chong et al. 2005). *id2* expression has been detected in migrating NCCs between 20-24 hpf, a period when fate segregation takes place (Chong et al. 2005). Diverse functions of the *id2* paralogue, *id2a*, have been suggested in several tissues during zebrafish development, including neurogenesis in the retina (Uribe and Gross 2010) and hepatoblast proliferation (Khaliq et al. 2015). Frederico Rodrigues, a PhD student in the Kelsh laboratory, showed that *id2a* is expressed in premigratory and migrating tNC populations, as well as in mature iridophores. Furthermore, morpholino-mediated knockdown of *id2a* completely abolished iridophore development, although no observable effects on other pigment cell lineages were noted (F. Rodrigues, PhD thesis). Lack of *id2a* mutant lines to date prevent directly testing the gene's position in the iridophore GRN, therefore *id2a* was not included in the preliminary network described in chapter 3. Despite its initial exclusion, id2a was clearly a prime candidate for if our studies subsequently required an iridophore-specific repressor.

Forkhead box D3 (*foxd3*) is a winged helix domain TF, which has been suggested to promote iridophore development. Several studies in zebrafish and in other species have strongly suggested that Foxd3 is a key repressor of Mitfa during early stages of NC development (Kos et al. 2001; Thomas and Erickson 2009; Curran et al. 2010). Furthermore, the melanoiridoblast fate restriction model (Curran et al. 2010) proposed that Foxd3 represses *mitfa* expression, thus aiding specification of iridoblasts at the

expense of melanoblasts. However, the role of Foxd3 in NC development is multifaceted. Data from different model organisms, including zebrafish, support a role of Foxd3 in NC establishment (refer to section 1.3.2), delamination and migration processes (Cheung et al. 2005; M. S. Simões-Costa et al. 2012). More generally, Foxd3 has been assigned an important role in embryonic stem cells (ESCs), in which it is thought to inhibit apoptosis and differentiation while promoting self-renewal (Y. Liu and Labosky 2008; Zhu et al. 2014).

Recent studies identified Fox family genes as pioneer TFs, functioning to displace nucleosomes and recruit chromatin remodellers at enhancer sites, thus decompacting chromatin and rendering these enhancers poised to respond to environmental cues (Zaret and Carroll 2011; Adam et al. 2015). Pioneer TFs thus regulate developmental competence and function during several stages to mediate acquisition or maintenance of diverse fates. A known pioneer TF from the Fox family, FOXA1, has been implicated in poising enhancer elements to respond to signals, in particular to promote pancreatic development (Cirillo et al. 2002; Sérandour et al. 2011; A. Wang et al. 2015). Furthermore, FoxE and FoxO family transcriptional regulators have been shown to act as chromatin remodellers to allow transcriptional activation in different contexts (Cuesta, Zaret, and Santisteban 2007; Hatta and Cirillo 2007).

In line with it acting as a pioneer TF, FOXD3 was found to mediate dynamic developmental transitions of embryonic stem cells (ESCs) and epiblast cells (EpiCs) (Krishnakumar et al. 2016). FOXD3 was suggested to prime stem cell fate choice by binding to enhancer elements, prior to any other transcriptional regulator, recruiting the chromatin remodeller BRG1 and removing nucleosomes. At the same time, FOXD3 maintained its targets in a transcriptionally inactive state by recruiting HDAC1/2 on enhancer elements (Krishnakumar et al. 2016), which is particularly interesting in the context of zebrafish melanophore development, where *hdac1* has been shown to repress *sox10* in an *mitfa*-dependent manner. In a different experimental context, a different set of enhancers was targeted by murine Foxd3, which under those circumstances acted as a mediator of the transition between naive and primed pluripotency (Respuela et al. 2016). Importantly, in both cases the role of Foxd3 as a pioneer TF functioning in stem cell development was supported.

The role of Foxd3 in priming pluripotent cells to respond to microenvironment cues is likely to prove important in deciphering its function during NC development, specifically when considering pigment cell specification. Compelling evidence suggests that Foxd3 suppresses the melanocyte lineage gene Mitf. Likely direct regulation via promoter binding was suggested in zebrafish (Curran, Raible, and Lister 2009), or indirect repression through sequestration of the Mitf co-activator, Pax3, in avian embryos

(Thomas and Erickson 2009). The former interaction was proposed to bias progenitors to adopt iridophore rather than melanophore fate (Curran et al. 2010), while the latter to promote gliogenesis over melanogenesis *in vitro* (Thomas and Erickson 2009).

In zebrafish, the function of *foxd3* in pigment cell development remains a subject of controversy. A role of *foxd3* in iridophore development was supported by data obtained using a *foxd3*:GFP transgenic line, which suggested that the *foxd3* promoter is active in mature iridophores (Curran, Raible, and Lister 2009). Nevertheless, a previous study reported that only a proportion of iridophores expressed *foxd3* (James A Lister et al. 2006), while sophisticated characterisation of the *foxd3* expression pattern in the developing embryo using transgenic lines did not reveal *foxd3* expression in the iridophore lineage (Kwak et al. 2013). In embryos with induced loss of *foxd3* function, either through morpholino technology, or using different *foxd3* mutant lines (discussed in section 5.1), it was shown that iridophores along the dorsal but not the ventral stripe are reduced (James A Lister et al. 2006; Montero-Balaguer et al. 2006; Stewart et al. 2006). In addition, a reduction was evident in *foxd3;mitfa* double mutants compared to *mitfa* single mutants (Curran et al. 2010). Recently, it was reported that disruption of the *foxd3* coding sequence, prior to the winged helix domain, via insertion of a fluorescent marker results in noticeable increase in melanophore numbers, when in homozygosis (Hochgreb-Hägele and Bronner 2013). These results support the proposed melano-iridoblast fate switch model (Curran et al. 2010) and argue for including *foxd3* in the iridophore GRN.

In contrast to what would be expected according to the fate switch model, melanophore numbers showed no increase in *foxd3* mutant embryos compared to WT siblings (James A Lister et al. 2006; Montero-Balaguer et al. 2006; Stewart et al. 2006; Cooper, Linbo, and Raible 2009), thereby challenging the view that *foxd3* represses *mitfa*, thus promoting iridophore specification. Interestingly, loss of function studies have further shown that specification of NC derivatives is defective in *foxd3* mutants. Notably, instead of *mitfa* appearing derepressed at early stages of NC migration (approximately 24 hpf), its expression was delayed (James A Lister et al. 2006; Montero-Balaguer et al. 2006; Stewart et al. 2006). Taking into account the early roles of *foxd3* in NC development (refer to section 1.3.2) and its proposed role as a pioneer factor, it was concluded that the roles of *foxd3* in chromatophore development need to be addressed more rigorously before the gene is incorporated in the GRN.

1.5.5. Differentiation genes

In terms of differentiation genes which could function as early markers of the iridophore lineage, similar to *dct* in the melanophore lineage, *pnp4a* is thus far the favoured

candidate (Curran et al. 2010). *pnp4a* is the zebrafish orthologue of the mammalian Pnp4, a purine nucleoside phosphorylase functioning in the purine metabolic pathway (Moriwaki, Yamamoto, and Higashino 1999). Expression on *pnp4a* was reported from as early as 20 hpf in a small subset of cNCCs and, as development progressed, was shown to mark cells residing in iridoblast and mature iridophore positions (Curran et al. 2010).

More enzyme-coding genes have been reported which are likely expressed in iridophores. For example, *gart* and *paics* mutants were found to harbour defects in all embryonic chromatophore lineages, most severely affecting iridophores and xanthophores (Ng et al. 2009). Both *gart* and *paics* code for members of the inosine monophosphate (IMP). *atic* is another gene of interest. Its product catalyzes the ninth and tenth reactions of IMP synthesis, and its expression was found to be enriched in isolated iridophores (Marie et al. 2004; Higdon, Mitra, and Johnson 2013). Isolation of iridophores from live embryos based on their physical properties revealed an extensive list of enzyme-coding genes (Higdon, Mitra, and Johnson 2013), which will not be discussed further here.

1.6. Objectives

The gene regulatory interactions underlying NC establishment, as well as specification of NC derivatives such as PNS components, craniofacial derivatives and melanophores have been well established. The aim of this project was to construct the GRN governing iridophore development, which had not been previously investigated, using an iterative approach of mathematical modelling and experimental validation. Such an approach was successfully employed in the past to generate the melanophore GRN (Greenhill et al. 2011). This study firstly aimed to identify the interactions between candidate genes functioning in iridophore development. To this end, classical genetics approaches revealed several novel, as well as conserved amongst NC derivatives, regulatory features that comprised the preliminary network guiding iridophore fate choice.

Secondly, this work aimed to harness the power of mathematical modelling to improve and expand the determined preliminary GRN. Iterations of mathematical modelling, computational simulations and experimental evaluation of modelling predictions led to the identification of overlooked interactions and novel uncharacterised factors. Furthermore, building on the Greenhill et al. modelling approach, this work aimed to identify physiologically relevant parameter values for the mathematical model and to rigorously investigate the parameter space via newly generated algorithms. Finally, the study seeked to investigate the role of *foxd3* in pigment cell development in an attempt to reconcile the controversy in the published literature.

2

Materials and methods

2.1. Zebrafish methods

2.1.1. Zebrafish handling and husbandry

All fish strains used have been grown to adulthood in the fish facility at the University of Bath. Handling of both adults and embryos was according to instructions published in the "Zebrafish Book" (Westerfield 2007) and abiding by Home Office regulations.

Crosses were set up in the evening as outlined in section 2.7 of the Zebrafish Book ("Natural crosses in false bottom containers"). Embryos were collected the following day, rinsed with water from the system and immediately transferred to Petri dishes (round, 90 mm) containing embryo medium. Petri dishes were then placed in an incubator set to 28.5 ºC. Unfertilised, dead or abnormal embryos were removed and siblings were staged according to Kimmel et al. (Kimmel et al. 1995). In order to prevent retarded or abnormal growth, the density of embryos did not exceed 90 individuals per Petri dish.

Depending on which stage was required for each experiment, cleaned and staged zebrafish embryos were either placed back in the 28.5 ºC incubator, or transferred to designated incubators set to 23 ºC or 33 ºC. At these temperatures, the growth rate of zebrafish embryos is delayed or increased respectively (Kimmel et al. 1995).

To prevent melanisation, embryos were treated with 0.003% 1-phenyl-2-thiourea (PTU; Sigma-Aldrich, Cat# P7629), diluted in embryo medium prepared without the addition of methylene blue. To allow thorough examination, live mounting and imaging, embryos were anaesthetised using a 0.4% solution of ethyl 3-aminobenzoate methanesulfonate salt (Tricaine; Sigma-Aldrich, Cat# A5040).

Live embryos were manually dechorionated using a pair of forceps from 24 hpf onwards. Embryos were fixed using a solution of 4% paraformaldehyde (PFA; Sigma-Aldrich, Cat# 158127). For further details on fixation protocols refer to sections 2.5.2 and 2.6.1.

Embryos were euthanised by adding several drops of 1% phenoxyethanol to the Petri dishes containing embryo medium.

2.1.2. Maintenance and introduction of new fish lines

Fish lines were regularly crossed (every approximately 12 months) to obtain the next generation. WIK and AB wild-type (WT) lines were incrossed while fish carrying mutations were alternatively incrossed to their siblings and outcrossed to WT fish to maintain a degree of genetic diversity. To further increase genetic heterogeneity, embryos derived from different crosses were mixed prior to entering the main system. Upon entering the permanent aquarium of our fish facility, all fish were registered in an internal zebrafish database according to Home Office guidelines.

New fish strains were obtained from collaborating institutions either by shipping adult zebrafish that are identified mutation carriers, or by obtaining embryos produced in the respective institution by incrossing identified carriers. In the former case, adults were kept and crossed in a designated quarantine room. Eggs that were either shipped or collected in the quarantine room were bleached twice for 5 min with 0.003% hypochlorite solution, followed by rinsing in system water from the main facility prior to entering the main fish facility.

2.1.3. Fish strains used

The following zebrafish WT and mutant strains were used in the course of this project:

WT:

1) **AB** (http://zfin.org/action/genotype/view/ZDB-GENO-960809-7).
2) **WIK** (Rauch, Granato, and Haffter 1997).

Mutants:

1) ***sox10^{t3}***. Strain reported by Kelsh et al. (1996). Genomic DNA sequencing of *sox10^{t3}* homozygous mutants identified an insertion of 1,397 kb upstream of the *sox10* high-mobility group box (HMG) DNA-binding domain. The inserted sequence is homologous to that of an identified transposon and results in the introduction of eight amino acids prior to a premature stop codon (Dutton et al. 2001).
2) ***ltk^{ty82}***. An ENU-induced single nucleotide change (R N Kelsh et al. 1996) generated a premature stop codon at amino acid (a.a). 785 (Lopes et al. 2008). The kinase functional domain, beginning from a.a. 909, is therefore deleted.
3) ***mitfa^{w2}***. An ENU-induced single nucleotide change generated a premature stop codon (a.a. 113) (J A Lister et al. 1999). The important functional domains of this

protein as listed in Ensembl ('domains and features' tool) begin from a.a. 285, and are therefore deleted.

4) ***foxd3***^{sa20726}. The line was generated via ENU-mutagenesis in the duration of the Zebrafish Mutation Project in the Sanger institute, UK (2016). The mutant harbours a nonsense mutation in the first exon (a.a. 41 of the protein). According to Ensembl the fork head domain begins from a.a 94. (http://www.sanger.ac.uk/sanger/Zebrafish_Zmpgene/ENSDARG00000021032#sa174).
5) ***tfec***$^{-/-}$. Generated during this project by clustered regularly interspaced short palindromic repeats (CRISPR)/ CRISPR-associated 9 (Cas9) mutagenesis, in collaboration with Dr James Lister (Virginia Commonwealth University).

2.1.4. Live imaging and counting iridophores

To mount zebrafish embryos for live imaging, 1% low melting point agarose (Invitrogen; Cat# 16520050) in filtered embryo medium (prepared without methylene blue) was heated in a water bath set at 70 ºC. Once the agarose was fully dissolved, 1 ml aliquots were prepared to allow efficient cooling of the solution immediately prior to mounting the embryos. Embryos were anaesthetised using 0.4% Tricaine solution and individually transferred on a mounting dish with glass coverslip bottom, where they were appropriately oriented. A drop of agarose cooled to approximately 40 ºC was placed on the embryo to be mounted. While it solidified, the embryo was being positioned using a pair of fine mounting needles.

Live imaging of iridophores under incident light was carried out using the Zeiss Axio Zoom.V16 fluorescence stereo zoom microscope (camera: Zeiss Axio Cam MRm). Brightfield images were taken using a Zeiss Imager 2 microscope (camera: Axiocam506). Following imaging, the embryo was rescued for further growth or alternative downstream applications by carefully breaking the surrounding agarose and placing it back in embryo medium.

Iridophores were counted using standard dissecting microscopes (Leica MZ75) and incident light. Embryos were PTU-treated to inhibit melanisation and allow more accurate scoring of iridophores directly adjacent to melanophores. Iridophores were easily identified due to their reflective properties and regularly spaced positioning in the dorsal and ventral stripes. For statistical analysis purposes, iridophores of the dorsal and ventral stripes only were counted. Iridophores of the lateral patches and of the eyes were not scored. When investigating potential iridophore phenotypes, iridophores of the whole embryo were analysed.

Iridophores tend to form clusters in which several cells become indistinguishable from each other. When counting, single iridophores were considered as any reflective signal that could be considered separate from the adjacent ones based on its round shape and relatively small size. Clusters of iridophores were determined based on their elongated, or otherwise unusual shape, larger size and often increased intensity of light reflection, when compared to single spots. Based on the shape and size of each cluster, the likely number of constituent single iridophores was estimated.

2.1.5. Solutions

Embryo medium

50x EM:

14.69 g/Lt NaCl (Molar mass: 58.44 g/mol, hence 251.4 mM)

0.63 g/Lt KCl (Molar mass: 74.5513 g/mol, hence 8.45 mM)

2.43 g/Lt $CaCl_2$ (Molar mass: 110.98 g/mol, hence 21.9 mM)

4.06 g/Lt $MgSO_4$ (Molar mass: 120.366 g/mol, hence 33.73 mM)

1x embryo medium:

20 ml stock

1 ml Methylene blue (10^{-4} %)

RO water to 1 Lt

Final concentration of salts:
Sodium Chloride: 5 mM

Potassium Chloride: 0.17mM

Calcium chloride: 0.44 mM

Magnesium sulfate: 0.67 mM

Phenoxyethanol

2-Phenoxyethanol, 99% (Fisher scientific; Cat# 10511571) was diluted 1:100 in water from the fish facility system.

PTU

Stock solution

0.03 g PTU powder dissolved in 100 ml embryo medium, prepared without methylene blue.

Working solution

1:10 dilution of the stock solution in EM.

Tricaine

400 mg tricaine powder
97.9 ml DD water
~2.1 ml 1 M Tris (pH 9).

Adjust pH to ~7.

2.2. Gene expression manipulation

2.2.1. Morpholino-mediated knockdown

For the morpholino microinjection, very thin needles were generated by melting 100 mm thin wall glass capillaries (1 mm outer/0.75 mm inner diameter) (World Precision instruments, Inc.) at 63 ºC using a needle puller (PC-10 puller; Narishige Ltd.). The tip of the needle was broken with a pair of forceps prior to injection, to allow the appropriate amount of liquid to pass through. The needle was backfilled with mineral oil and attached to a Drummond Nanoject II Auto-Nanoliter Injector (Drummond Scientific Co.). The tip of the needle was filled with the solution to be injected.

To inject, 90 mm Petri dishes were filled in advance with 2% agarose in embryo medium and a glass slide (VWR; 6310114) was placed on top of the solidified layer. Plates were maintained at 4 ºC for multiple uses. Embryos were collected (as described in section 2.1.1.) and positioned along the side of the glass slide. The embryo medium was drained and the embryos were injected at 1-2 cell stage (10-45 min post fertilisation) by piercing the chorion and the side of the yolk sac. Injections were carried out under a standard dissecting microscope (Leica MZ6). Unless stated otherwise, approximately 5 nl per embryo were deposited into the yolk sac, in close proximity to the cell. Morpholino molecules are able to cross the yolk sac-cell boundary via cytoplasmic streaming.

The *tfec* morpholino used in this work was designed by the J. Lister research group (Virginia Commonwealth University) and synthesised by Gene Tools, LLC. The molecule targeted the exon 5-intron 5 boundary, thus preventing mRNA splicing. The sequence is as follows:

E5I5: 5'-AGCCTAAAAACCACTTACTTAGATT-3'

The lyophilised morpholino was appropriately diluted in Milli-Q water to a final concentration of 1 mM and stored at room temperature to avoid precipitation. On the morning of injection, the stock was heated to 65 ºC for 5 min to dissolve any precipitation and appropriate dilutions were prepared. For maximum efficiency, the final injected volume per embryo was 4.6 ng-6 ng. To help visualise the injected solution, a final concentration of 0.025% phenol red dye was added.

2.2.2. CRISPR/Cas9-mediated targeted mutagenesis

The CRISPR/Cas9 system for site-directed mutagenesis takes advantage of the prokaryotic defence mechanism against bacteriophage infection. The principle is that a guide RNA (gRNA) molecule, designed to recognise a particular DNA sequence, forms a complex with the Cas9 endonuclease and guides it to the target site on the genomic DNA. Once the DNA sequence is bound, a typical Cas9 protein creates a double-

stranded break, which can be repaired by the cellular machinery in two distinct ways: by non-homologous end-joining (NHEJ) and by homology-directly repair (HDR). The former is an inherently error-prone mechanism, which can induce random insertions or deletions in the target sequence, thus often leading to frameshift mutations and to premature stop codons. The latter relies upon recombination with a template donor sequence, thus leading to specified alterations to the sequence, or to insertion of desired sequences in the genome at the position of the target site. Homology-directed repair will not be discussed further.

The gRNA comprises a sequence of approximately 20 nucleotides at the 5' end, which are complementary to the genomic target site, followed by three nucleotides complementary to the genomic protospacer adjacent motif (PAM) sequence, a scaffold region and a terminator signal. The mechanism of action of the CRISPR/Cas9 complex is described in figure 2.1.

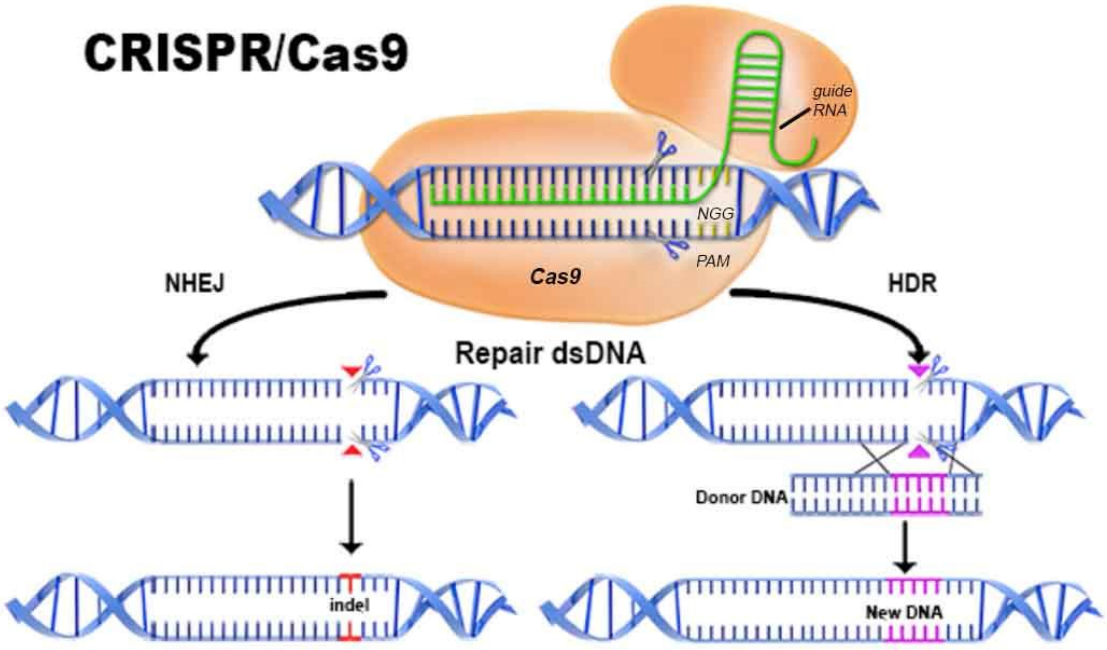


Figure 2.1. The mechanism of CRISPR/Cas9 site-directed mutagenesis. The CRISPR gRNA (green) and the Cas9 protein (orange) form a complex. The 5' end of the gRNA binds its complementary sequence in the genomic DNA, which is followed by a canonical PAM site (5'- NGG -3'). DNA repair mechanisms either facilitate NHEJ and random insertions or deletions ("indels") are induced, or HDR, in which case recombination with a donor DNA fragment takes place. gRNA, guide RNA; NHEJ, non-homologous end-joining; HDR, homology-directed repair.

RNA preparation

CRISPR/Cas9-mediated mutagenesis was employed to disrupt two genes: *id2a* and *tfec*. To knock-out *id2a*, two different CRISPR gRNA sequences (referred to as *id2a* CRISPR1, *id2a* CRISPR2), cloned into DR274 plasmid vectors, were obtained from the Sanger Institute (for the plasmid map refer to Appendix I.1). The target sequences were the following:

id2a CRISPR1: 5'- GGCTGAGAGGATCGTCCACGGGG -3'

id2a CRISPR2: 5'- GGTGATTCACCTGTAGTGAGAGG -3'

The Cas9 sequence, cloned into pCS2 plasmid vector, was also provided by the Sanger Institute (for the plasmid map refer to Appendix I.2).

Bacteria transformed with the plasmids were grown and the vectors were purified via midiprep (see sections 2.4.2-2.4.4). To generate CRISPR gRNA, 5 µg of plasmid was digested with *DraI* (Thermo Scientific; Cat# ER0221). The desired DNA fragments were isolated by agarose gel extraction (refer to section 2.3.5) and used as templates for *in vitro* transcription (IVT; refer to section 2.3.10 (2)). CRISPR gRNA was diluted to 100 ng/µl and aliquots were stored at -80 ºC. To generate Cas9 RNA, 5 µg of vector were linearised with *NotI* (Promega; Cat# R6431) and the desired product was gel-purified and used as template for IVT (refer to section 2.3.10 (3)). Cas9 mRNA 500 ng/µl aliquots were stored at -80 ºC.

The CRISPR sequence for disrupting *tfec* was designed by our collaborator, Dr J Lister, targeted exon 7 of the gene and was as follows:

5'- GA**CGATCCTCAAGGCCTCGG** -3'

A previously described cloning-free protocol (Talbot and Amacher 2014) was implemented to generate gRNA. Instead of ordering a single guide-oligo sequence, two shorter oligonucleotides were ordered from Eurofins MWG:

Short Guide:

5'-
GCGTAATACGACTCACTATAGG**CGATCCTCAAGGCCTCGG**GTTTTAGAGCTAGAA
ATAGC -3'

Guide Constant:

5'-
AAAGCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCTTATTTTAA
CTTGCTATTTCTAGCTCTAAAAC -3'

gRNA primer 1:

5'- GCGTAATACGACTCACTATAG -3'

gRNA primer 2:

5'- AAAGCACCGACTCGGTGCCAC -3'

The following elements are indicated above: the T7 RNA Polymerase promoter region, the **guide sequence**, the Cas9 binding scaffold and the region of overlap between the short guide and the guide constant oligonucleotides. The PCR reaction was set up by mixing the following reagents: 10 µl of 10X Thermopol Buffer (NEB; M0267S), 2.5 µl of 10 µM gRNA primer mix, 2 µl of 1 µM short guide oligo, 2 µl of 1 µM guide constant oligo, 3 µl DMSO, 2 µl 10 mM dNTPs, 1.5 µl Taq DNA Polymerase (NEB; M0267S) and sterile water to 100 µl. The reaction was divided into two 50 µl reactions which were run using the following conditions: initial denaturation at 95 ºC for 1 min, then 40 cycles of 95 ºC for 15 sec, 60 ºC for 30 sec and 72 ºC for 20 sec, followed by a final elongation step at 72 ºC for 5 min. The product was gel-purified and sequenced using the T7 promoter, before being used as template for IVT, as described in section 2.3.10 (2).

Injections

To optimise mutagenesis efficiency and survival rates, 1 nl of RNA was delivered inside the flat cell of fertilised embryos. Solution to be injected was prepared by thawing CRISPR guide RNA and Cas9 mRNA on ice and mixing 100 ng of CRISPR gRNA with 250 ng Cas9.

Injections were performed under a standard dissecting microscope using a Pneumatic Picopump (World Precision Instruments) and a MN-153 micromanipulator (Narishige). Borosilicate glass capillaries (1 mm outer/0.78 mm inner diameter) with filament (Clark Electromedical Instruments; Cat# GC100TF10) were melted at 63 ºC using a needle puller (PC-10 puller; Narishige Ltd.). The injection needle was back-filled with mRNA injection solution. Prior to aligning embryos on the injection dish (refer to section 2.2.1), the injected volume was measured using a S78 Reflected Light Stage Micrometer (1 mm/ 0.01 mm divisions) (Pyser; Cat# 02B00421), and adjusted to deliver 1 nl per injection.

Mutagenesis screening

In the case of *tfec*, the iridophore patterns of injected embryos were screened to assess mutagenesis efficiency. For *id2a*, where phenotypes were not obvious, screening of the F0 was done by sequencing. First, approximately 8 injected embryos between 24-48 hpf were placed into microfuge tubes and total genomic DNA was extracted via the NaOH

method (refer to section 2.3.5). The genomic DNA was used as a template to amplify the fragment of interest (refer to Appendix table II.4) using DreamTaq DNA Polymerase (Thermo Scientific; Cat# EP0702). The generated amplicons were ligated into the pGEM T Easy Vector System (Promega; Cat# A1360) (refer to section 2.4.1, and to Appendix section I.3) and were transformed into competent cells (refer to section 2.4.2). 16 grown single colonies were selected from the LP/agar plate, miniprep cultures were grown and the plasmids were purified, as described in section 2.4.4. Vectors were sent for sequencing from the T7 promoter, and sequences were analysed as described in section 2.3.6.

To genotype mutation carriers from F1 and F2 generations, incrossing and phenotypically screening embryos as well as High Resolution Melt Analysis (HRMA; refer to section 2.3.7) were used.

2.2.3. Overexpression studies

2.2.3.1. Construct preparation

mRNA to be injected was mainly prepared from plasmid templates. Circular plasmids were digested with appropriate restriction enzymes (refer to table 2.1) prior to mRNA synthesis using mMESSAGE mMACHINE® SP6 Transcription Kit (Ambion; Cat# AM1340) (refer to section 2.3.10 (3)). Restriction digests were assembled according to manufacturer's instructions and incubated at 37 ºC for 4 h. To ensure complete linearisation, a small amount of product was assessed using agarose gel electrophoresis. The rest of the reaction was then purified (refer to section 2.3.5), the IVT reaction was assembled and the product was purified according to the protocol described in section 2.3.10 (3).

To generate *tfec* mRNA, trunks and tails from 11 live WIK embryos at 3 days post-fertilisation (dpf) were isolated and immediately placed on ice. Total RNA was extracted using TRI reagent (refer to section 2.3.8) and DNaseI treatment was performed. For cDNA synthesis, the SuperScript® III First-Strand Synthesis SuperMix (Invitrogen; Cat# 18080400) was used according to manufacturer's instructions. 1 µg total trunk RNA was added as template and the recommended amount of supplied oligoDT_{20} was used.

To amplify the coding sequence of tfec, the following primer set was used:

FW: 5'- AGC GAG ATC CTC CTG CTT CG -3'

RV: 5'- ATT CTG AGA GTG CGG TCC AG -3'

The expected PCR product was 1.3 kb long and corresponded to one of the two *tfec* transcripts, as documented in Ensembl (ENSDART00000164766.1). Amplification was carried out using KOD Hot Start High-Fidelity DNA polymerase (Novagen, Cat# 71086)

according to the protocol described in section 2.5.1. The annealing temperature for the reaction was set as 56.4 ºC.

To attach the T7 promoter on the 5' end of the amplified cDNA sequence, a second PCR reaction was performed using KOD Hot Start High-Fidelity DNA polymerase (Novagen, Cat# 71086). The amplicon generated in the previous reaction was diluted 1:50 and 1 µl of the dilution was used as template in the second reaction, which again was performed as described in section 2.5.1, with an annealing temperature of 56.4 ºC. The reverse primer used in this reaction was the same as previously (RV: 5'- ATT CTG AGA GTG CGG TCC AG -3'). The forward primer was designed as a fusion of the T7 promoter sequence, as described in Ambion documentation, and the previously used forward primer:

T7/FW:5'- TAA TAC GAC TCA CTA TAG GGA GA AGC GAG ATC CTC CTG CTT CG

T7 promoter sequence FW primer sequence

The entire PCR reaction was run on an agarose gel and the band corresponding to the expected sequence length was extracted (refer to section 2.3.5). IVT was set up as described in section 2.3.10 (4), using 90 ng of purified amplicon as template for mRNA synthesis. An RNA gel (refer to section 2.3.11) was run to compare *tfec* mRNA lacking poly(A) tail and mRNA subjected to the poly(A) tailing reaction.

Table 2.1. Identity and handling of plasmid vectors used for overexpression mRNA preparation.

Gene name	Plasmid backbone	Origin/ Reference	Sequence length	Linearisation enzyme	IVT enzyme
***sox10*^{WT}** ***sox10*^{m618}**	pCSHSP	(Dutton et al. 2001)	1.6 kb	*Asp718I* (Roche; Cat# 10814245001)	SP6 RNA Polymerase
***mitfa*^{WT}** ***mitfa*^{b692}**	pCS2+	Lister group	1.3 kb	NotI (Promega; Cat# R6431)	
npm/ltk ***npm/npm-ltk*****(DK)**	pCS2+	Kelsh group	2.4 kb		
XLT.GFP$_{LT}$	pCS2+	Moon group (Carney et al. 2006)	c. 700 bp		

All RNA generated for injection was immediately diluted to 50 ng/µl using molecular biology grade water (Merck Millipore; Cat# H20MB0501). Subsequently, 5, 10 and 20 µl aliquots were prepared in 200 µl sterile tubes, for use in thermal cyclers, and stored at -80 ºC until required for injections.

2.2.3.2. Microinjection and sample preparation

To co-overexpress genes of interest, relevant mRNAs were mixed and injected into the yolk sac of single-cell stage embryos, and allowed to cross the yolk sac-cell boundary via cytoplasmic streaming. mRNA aliquots stored at -80 ºC were defrosted on ice. Equal volumes of pairs of different mRNAs were mixed to generate the injection solution, in which the concentration of each mRNA is 25 ng/µl. Preparation of the needle and injection apparatus, as well as the injection procedure, was as described in section 2.2.2.

Injected embryos were placed at 28.5 ºC and monitored. Total RNA was extracted at 6 hpf and 10 hpf (refer to section 2.3.8) from pools of 8 embryos per microfuge tube, and was stored at -20 ºC. RNA was treated to remove genomic DNA contaminants (refer to section 2.3.8) and cDNA synthesis was performed (refer to section 2.3.9) using 1 µg of total RNA template per reaction. cDNA was stored at -20 ºC until needed.

2.2.3.3. Quantitative real time PCR (q-RT PCR)

To quantitate levels of gene expression following mRNA overexpression, q-RT PCR was performed using Fast SYBR Green Master Mix (Applied Biosystems; Cat# 4385617). Reactions were assembled in MicroAmp® Fast Optical 96-Well Reaction Plates (Applied Biosystems; Cat# 4346907) and run using the StepOnePlus™ Real-Time PCR System (Thermo Scientific; Cat# 4376600).

Each reaction was assembled as follows: 8 µl of 2X Fast SYBR Green Master Mix, 0.8 µl of 5 µM primer mix (or ready-made *rlp13*/*dct* primer mix from Primerdesign Ltd.), 1 µl cDNA (diluted 1:4 or 1:6) and 6.2 µl molecular biology grade water were added to each well. The plate was tightly sealed, briefly centrifuged and placed in the StepOnePlus machine. The running protocol was as follows: initial denaturation at 95 ºC for 20 sec, followed by 40 cycles of denaturation at 95 ºC for 3 sec and annealing/elongation at 60 ºC for 30 sec. To generate the melt curve, the temperature was set to increase at 1.5 ºC per second increments from 60 ºC to 95 ºC.

Results were visualised using the StepOnePlus software and exported as an Excel file. To estimate relative expression change (fold activation or repression), Double Delta C_T analysis was used. The following steps were followed using the exported excel spreadsheet:

1) Normalisation of mean C_T values calculated for each target gene to those of the housekeeping gene:

$$\Delta C_T = C_T^{target} - C_T^{housekeeping}$$

2) Averaging ΔC_T values of pairs of biological replicates for each condition.
3) Calculating the $\Delta\Delta C_T$ value, by subtracting ΔC_T [treated-untreated]. In this case the 'untreated' sample corresponded to GFP-injected embryos.

$$\Delta\Delta C_T = \Delta C_T^{treated} - \Delta C_T^{untreated}$$

4) Deriving the fold change

$$Fold\ change = 2^{-\Delta\Delta C_T}$$

Fold change values for each condition, relative to untreated controls were plotted in Excel.

2.3. Molecular methods

2.3.1. Genomic DNA isolation

To isolate genomic DNA from live zebrafish embryos, two different protocols were used. The first was based on sodium hydroxide-mediated lysing and achieved crude extraction of genomic DNA to be used for routine genotyping. A single embryo or group of embryos was placed in an microfuge tube, the embryo medium was removed and replaced with 50 mM NaOH (Fisher Scientific; S/4920/53). For a single embryo, 50 µl NaOH was added. To lyse more embryos the volume of added NaOH was scaled up, to a maximum of 500 µl per sample. Two to three openings were made on the lid of the microfuge tube using a syringe needle (27G) to prevent the tube from bursting. Embryos were incubated at 95 ºC for 5-10 min or until completely dissolved. The samples were then chilled on ice and neutralised by adding Tris HCl pH 8.0 to a final concentration of 0.1 mM. The solution was mixed well and centrifuged at 13,000 rpm for 5 min to pellet debris. The supernatant was transferred to a sterile microfuge tube and stored at -20 ºC until needed.

Higher purity genomic DNA for quantitative real time PCR (q-PCR) and High Resolution Melt Analysis (HRMA) assays was obtained using the KAPA Express Extract kit (KAPA Biosystems; Cat# KK7103). The kit contains 10X concentrated buffer solution and a novel thermostable protease. The extraction solution was prepared according to manufacturer's instructions. To process single embryos, the samples were transferred to thin-walled 100 µl or 200 µl tubes, all embryo medium was removed and 30 µl of extraction solution was applied to each tube. The samples were placed in a thermal

cycler and incubated at 60 ºC for 10 min, then at 95 ºC for 5 min, as per manufacturer's instructions. Subsequently, the samples were briefly vortexed, centrifuged for 1 min to pellet the debris and either stored at 4 ºC for up to one month or long term at -20 ºC. To use extracted DNA as template for HRMA, the concentration of genomic DNA in each sample was quantified and 8 ng/µl dilutions were prepared.

Genomic DNA was also extracted from fixed embryos which had been previously processed for *in situ* hybridisation (refer to section 2.5). Single embryos were transferred to thin-walled 200 µl tubes for use in thermal cyclers, and were washed with 50 µl sterile water to remove remaining glycerol. Water was then replaced with 30 µl of buffered proteinase K solution per embryo (refer to solutions, section 2.3.12). The tubes were incubated in a thermal cycler (G-Storm 482) at 50 ºC for 4 h. During that time they were briefly vortexed every 30-40 min. The enzyme was deactivated at 95 ºC for 10 min. Samples were centrifuged to pellet the debris and stored at 4 ºC to be used within a few days or at -20 ºC for long term storage.

To isolate genomic DNA from adult zebrafish, the newly established swabbing technique was employed (developed by Breacker C. & Tilley C. A., University of Leicester; Breacker & Tilley 2014). To isolate genomic DNA via swabbing, the fish of interest was caught with a net and placed on a flat wet sponge. The animal was gently secured in a lateral position using the thumb and middle finger to hold it ventrally and dorsally, respectively. Folds of the net, held gently in place by the forefinger, were used to cover the head of the fish, thus leaving the trunk and tail exposed. A regular cotton bud was used to swab in an anterio-posterior direction a maximum of five times along the flank between the anal and the caudal fin, as described by Breacker & Tilley. The cotton bud was placed in a microfuge tube on ice until all fish of interest could be swabbed.

This technique is advantageous compared to the previously used fin-clipping method, as it eliminates the need for anaesthesia, thus avoiding stressing and physically harming the animal while rendering the process faster. Additionally, it is not an invasive technique, therefore minimising the risk of pain, secondary infections and behavioural changes. Finally, it has the added benefit that it is not considered a licensed procedure.

To extract the genomic DNA, the tip of the cotton bud was detached using a pair of scissors and uniformly wetted with 60 µl of KAPA Express Extract solution, prepared according to manufacturer's instructions. The microfuge tube containing each soaked tip was sealed and placed in a water bath set to 60 ºC for 10 min, then the enzyme was inactivated by heating to 95 ºC for 5 min. The samples were returned to ice and the cotton buds were placed into sterile 0.5 ml tubes, of which the lids were removed and the bottom was pierced using a pair of forceps, creating a hole approximately 2 mm wide. The 0.5 ml tubes containing the cotton buds were placed inside the original 1.5 ml

microfuge tubes and the assembly was centrifuged for 5 min to isolate the aqueous solution containing the extracted DNA from the cotton bud. The samples were stored at -20 ºC until use.

For each of the techniques described above, concentration of extracted DNA was determined by using the NanoDrop 2000 UV-Vis Spectrophotometer (Thermo Scientific), using extraction buffer as a blank.

2.3.2. PCR-based genotyping

DreamTaq DNA Polymerase (Thermo Scientific; Cat# EP0702) was used to amplify DNA fragments between 100-1,000 bp, according to manufacturer's instructions. For amplification from more challenging or compromised genomic DNA samples (such as those where either highly crosslinked chromatin or contaminants were present owing to the nature of the samples or the extraction methodology, as well as either GC- or AT-rich templates), the KAPA2G Robust PCR kit was used (HotStart ReadyMix with dye; KAPA Biosystems; Cat# KK5704). All PCR reactions were run in the G-Storm 482 thermal cycler.

To genotype carriers of the *foxd3*^{sa20726} allele, the following reaction conditions were used: 5 µl 2X KAPA2G Robust Master Mix, 0.5 µl each of 10 µM forward and 10 µM reverse primer (Appendix table II.1), 0.5 µl 100% DMSO (Finnzymes; Cat# F515), 3 µl 10 mg/ml purified BSA (NEB; Product discontinued, replaced by Cat# B9000S) and 0.5 µl genomic DNA template were mixed for each 10 µl reaction. The cycling conditions were as follows: initial denaturation at 95 ºC for 3 min followed by 40 cycles of incubation for 15 sec at 95 ºC, 15 sec at 55 ºC and 60 sec at 72 ºC. The samples were stored in the thermal cycler (G-Storm 482) at 10 ºC. Use of BSA in the above protocol was initiated based on work showing that it enhances amplification from GC-rich templates in the presence of DMSO (Farell and Alexandre 2012).

Amplicons were then digested with *DdeI* enzyme. Following digestion, the WT amplicon (227 bp) was unaffected, whereas the mutant amplicon was digested producing two fragments (122 bp and 105 bp). For the restriction digest, the following reagents were added in each 10 µl PCR reaction: 2.5 µl 10X Buffer 3 (NEB; Cat# B7003S), 0.5 µl 10 U/µl *DdeI* (NEB; Cat# R0175S) and 12 µl sterile water, to a final volume of 25 µl. The reactions were incubated for 2-4 h at 37 ºC in a thermal cycler (G-Storm 482), and the enzyme was deactivated at 65 ºC for 20 min. Although the KAPA2G Robust Master Mix includes dye, to ensure proper sinking of the restriction digest reaction in the wells during agarose gel electrophoresis, an additional 3 µl of 6X Blue/Orange Loading Dye (Promega; Cat# G1881) were added to each sample.

To genotype carriers of the *ltk*^{ty82} mutant allele, the following reaction conditions were used: 5 µl 2X KAPA2G Robust Master Mix, 0.5 µl each of 10 µM forward and 10 µM reverse primer (Appendix table II.1), 0.5 µl genomic DNA template and 3.5 µl sterile water per 10 µl reaction. The cycling protocol, which was previously optimised in the group (Lopes et al. 2008), was carried out as follows: initial denaturation at 95 ºC for 3 min followed by 35 cycles of incubation for 30 sec at 94 ºC, 30 sec touchdown from 55–47 ºC and 30 sec at 72 ºC. A final extension of 10 min at 72 ºC was allowed.

As previously established (Lopes et al. 2008), amplicons were digested with *NheI* enzyme. WT amplicons were unaffected (490 bp), whereas mutant amplicons generated two fragments (360 bp and 130 bp). For the restriction digest, the following reagents were added in each 10 µl PCR reaction: 3 µl 10X Buffer B (Promega; Cat# R002A), 0.3 µl 10 mg/ml BSA (Promega; Cat# R3961), 0.7 µl 10 U/ µl NheI (Promega; Cat# R6501) and 16 µl sterile water for a final volume of 30 µl. The reactions were incubated for 4 h at 37 ºC in a thermal cycler (G-Storm 482) and the enzyme was deactivated at 65 ºC for 20 min. An additional 4 µl of 6X Blue/Orange Loading Dye (Promega; Cat# G1881) were added to each sample prior to agarose gel electrophoresis.

2.3.3. Agarose gel electrophoresis for DNA

For most applications, DNA and RNA products were routinely visualised using agarose gel electrophoresis. 1-3% (w/v) agarose gels were made either in 1X TAE or in 1X TBE buffer (refer to solutions, section 2.3.12) and with a final concentration of 1 µg/ml ethidium bromide. Samples containing 1X loading dye were loaded into the wells and voltage of 60 – 130 V was applied depending on the purpose of the experiment, concentration of the gel and type of nucleic acid. Typically, genotyping gels (3% agarose in TAE) were run at 80-90 V, whereas 1-2% gels loaded with DNA were run at 90-100 V. 1% RNA gels were run at higher voltages to reduce the length of running time and, thus, to avoid potential degradation of the RNA. Agarose gels were visualized and photographed under UV light using the ChemiDoc-It®^{2} Transilluminator Imager (UVP).

2.3.4. Acrylamide gel electrophoresis

DNA acrylamide gel electrophoresis was used to achieve a higher degree of separation and increased sensitivity for applications such as CRISPR/Cas9 mutagenesis screening and genotyping through restriction digestion.

For acrylamide gel electrophoresis, 12-15% gels were cast. First, the casting apparatus was assembled. Mini-PROTEAN Spacer Plates with 1.0 mm Integrated Spacers (Biorad; Cat# 1653311) and Mini-PROTEAN Short Plates (Biorad; Cat# 1653308) were carefully cleaned, sprayed with 70% ethanol (EtOH) and allowed to air dry. Short plates were placed on top of spacer plates, the edges were carefully aligned and the assembly was

fitted in Mini-PROTEAN Casting Frames (Biorad; Cat# 1653304). The casting chambers were assembled by fitting thermoplastic rubber gaskets (Biorad; Cat# 1653305) on Mini-PROTEAN Casting Stands (Biorad; Cat# 1653303) and securing the casting frames on the stands so that leakage of liquid acrylamide solution was prevented.

To make 30 ml of solution for casting four 15% gels, the following reagents were mixed: 8.8 ml of distilled water, 6 ml of 5X TBE (refer to solutions, section 2.3.12) and 15 ml of 30% acrylamide (AccuGel 29:1; National Diagnostics; Cat# EC-851). After thorough mixing, 210 µl of 10% (w/v) Ammonium Persulfate (APS) and 21 µl of Tetramethylethylenediamine (TEMED) were added to allow crosslinking to occur. The solution was mixed quickly and thoroughly by inverting, and a P1000 Gilson pipette was used to fill the space between the spacer plate and the short plate with the liquid solution. 10-well or 15-well Mini-PROTEAN Combs (Biorad; Cat# 1653359, #1653360) were fitted into the space on top of the plates and the gels were allowed to set for 1-2 h at room temperature. To store the gels for later running, paper towels were soaked with 1X TBE and wrapped around the plates to prevent the gels from drying. Cling film was then wrapped around the paper towels and the package was placed at 4 ºC.

To run the gels, a Mini-PROTEAN Tetra Electrode Assembly (Biorad; Cat# 1658037) and a Mini-PROTEAN Tetra Companion Running Module (Biorad; Cat #1658038) were used. The gels were fitted on the device, with the short plates facing inwards, to form a watertight seal. If necessary, a Mini Cell Buffer Dam (Biorad; Cat #1653130) was used to complete the set up. The assemblies were fitted in a buffer tank (Biorad; Cat# 1658040) and 1X TBE buffer was poured to fill the spaces between the inward-facing gels. The combs were removed and 1X TBE buffer was used to fill the tank to the suitable level, as designated on the tank. 5 µl of GeneRuler 50 bp DNA Ladder (Thermo Fisher; Cat# SM0371), prepared according to manufacturer's instructions, and the samples containing 1X Blue/Orange Loading Dye (Promega; Cat# G1881) were loaded. For 15% gels, approximately 80 V were applied (PowerPac Basic, Biorad) for 3-4 h, until the Bromophenol blue marker was nearing the edge of the gel.

The plates were then removed and the gels were exposed using a gel releaser (Biorad; Cat# 1653320). Gels were carefully transferred to a solution of 1 µg/µl ethidium bromide in 1X TBE and allowed to stain for approximately 45 min – 1 h before imaging using the ChemiDoc-It®^{2} Transilluminator Imager (UVP). The glass plates and combs were thoroughly cleaned to ensure no polymerised acrylamide remained. The remaining equipment was rinsed with water and allowed to dry.

2.3.5. DNA extraction and purification

DNA purification via gel extraction was carried out using the QIAquick gel extraction kit (QIAgen; Cat# 28704), according to the manufacturer's instructions. For direct purification of the amplicon from the PCR reaction the DNA Clean & Concentrator™-5 (Zymo Research; Cat# D4003) was used according to manufacturer's instructions. In both cases, purified DNA was eluted in Milli-Q water and quantified using the NanoDrop 2000 UV-Vis Spectrophotometer (Thermo Scientific).

For purification of linearised DNA following restriction digest, basic pH phenol-chloroform-isoamyl alcohol (Sigma Aldrich; Cat# 77617) was used. Milli-Q water was first added to the reaction to a final volume of 100 µl. The sample was transferred from a thin-wall 200 µl tube to a 1.5 ml microfuge tube and 100 µl of phenol-chloroform-isoamyl alcohol mixture were added. The mixture was vortexed for 15 sec and centrifuged at 13,000 rpm for 3 min to achieve phase separation. The upper phase, containing the DNA, was transferred to a fresh 1.5 ml microfuge tube and 2.5 volumes of 100% EtOH and 1/10 volume sodium acetate (3 M, pH 5.4) were added. DNA was precipitated at -20 ºC overnight. The following day, the sample was centrifuged at 13,000 rpm for 30 min at 4 ºC, the supernatant was decanted and the pellet was rinsed with 500 µl of ice cold 70% EtOH. The supernatant was again decanted and the pellet was air dried for approximately 5-10 min at room temperature, before being re-suspended in sterile water and quantified using the NanoDrop.

2.3.6. Sequencing

DNA sequencing was performed commercially by Eurofins MWG Operon, London, UK. Returned sequences were analysed using the Basic Local Alignment Search Tool (BLAST) at NCBI (http://blast.ncbi.nlm.nih.gov/Blast.cgi), the Multiple Sequence Alignment tool CLUSTALW2 (http://www.genome.jp/tools/clustalw/), or manually using the freely available software FinchTV (Geospiza). Sequence alignment was also conducted using freely available MEGA software.

2.3.7. Genotyping via High Resolution Melt Analysis (HRMA)

HRMA is a q-RT PCR-based approach used for identifying variations in nucleic acid sequence based on differences in the melt (dissociation) curves of amplicons. Each amplicon's precise melting temperature is determined by its base pair composition. Therefore, single nucleotide polymorphisms (SNPs) and larger insertions or deletions slightly affect the melting temperature of the corresponding amplicon. Using sensitive q-RT PCR instrumentation these changes can become detectable. Appropriate HRMA software analyses the data to detect deviations between unknown samples and reference control samples, which would suggest sequence alterations. In this study,

HRMA was employed as an alternative method of genotyping carriers of the *foxd3*^{sa20726} mutant allele. It was also used as a method to screen for changes in the genomic sequence as a result of CRISPR/Cas9 mutagenesis.

Firstly, appropriate primer sets were designed for each application (refer to Appendix table II.2). For HRMA, the amplicon length ranges between 80-200 bp. Primer sets that generate short amplicons (80-120 bp) were designed to more efficiently detect variations that cause small differences in melting temperatures (A→T or C→G single nucleotide polymorphisms (SNPs) produce differences in melting temperature of less than 0.2 ºC and 0.2-0.5 ºC respectively according to Applied Biosystems documentation). Modifications resulting from CRISPR-Cas9 mutagenesis are much more pronounced (typically several bases deleted or inserted), causing larger perturbations in melt curves and were accurately detected with amplicon sizes between 100-200 bp. To detect undetermined CRISPR/Cas9 modifications, primers were designed at least 20 bp on either side of the CRISPR recognition site, to prevent large deletions from affecting the primer binding site. Presence of natural sequence variants in the amplified region was also considered, as this may result in false positives or negatives. The amplicon sequence was screened for presence of naturally occurring SNPs using the "variation features" track on Ensembl (www.ensembl.org). Furthermore, to ensure that the amplicon's sequence was able to generate a smooth melt curve, with a single melt peak, the freely available web-based application uMelt (https://www.dna.utah.edu/umelt/umelt.html) was used.

To set up the qPCR reaction for HRMA, KAPA HRM FAST reagents were used (KAPA Biosystems; Cat# KK4201). The KAPA HRM FAST Master Mix contains EvaGreen saturating dye, which allows for increased sensitivity and effective discrimination between sequence variants. Reactions were set up either in MicroAmp® Fast Optical 96-Well Reaction Plate (Applied Biosystems; Cat# 4346907) or in MicroAmp® Fast 8-Tube Strips (Applied Biosystems; Cat# 4358293).

Reactions were set up as follows: 10 µl 2X KAPA HRM FAST Master Mix, 1.2 µl or 2 µl 25 mM $MgCl_2$ (final concentration 1.5 mM and 2.5 mM respectively), 0.8 µl 5 µM forward/reverse primer mix, 1 µl genomic DNA template (8 ng/µl) and molecular biology grade water to 20 µl were added per well. In order to avoid detectable contamination, all primers dilutions and qPCR master mixes were assembled using certified water for molecular biology (Merck Millipore; Cat# H20MB0501).

The reactions were run using the StepOnePlus™ Real-Time PCR System (Thermo Scientific; Cat# 4376600). Cycling conditions were as follows: initial denaturation at 95 ºC for 3 min, followed by 40 cycles of incubation at 95 ºC for 5 sec, (55 ºC – 60 ºC) for 20 sec and at 72 ºC for 10 sec. A continuous melt curve was generated, first by incubating

at 95 ºC for 15 sec, then by increasing the temperature from 60 ºC (1 min) to 95 ºC (15 sec) in 0.3 ºC /sec increments. Finally the temperature was dropped to 60 ºC for 15 sec.

Amplification and melt curve results were analysed using the StepOnePlus Software (Applied Biosystems). HRMA was performed using High Resolution Melting (HRM) Software V3.0.1 (Applied Biosystems) according to manufacturer's instructions. For each assay, 4-16 reference samples were set up. To detect CRISPR/Cas9-mediated mutagenesis, at least 8 WT reference samples were included. To genotype mutant alleles, at least 4 samples for each of known WT, heterozygous and homozygous mutants were included.

2.3.8. Total RNA isolation

For extraction and isolation of total RNA, TRI Reagent (Sigma-Aldrich; Cat# T9424) was used. The extraction protocol was based on the single step total RNA isolation (Chomczynski and Sacchi 1987). Up to 10 embryos were collected in 1.5 ml microfuge tubes, all embryo medium was removed and replaced with 200 µl of TRI reagent. The tissue was homogenised using micropestles and incubated at room temperature for 5-15 min. Then, 40 µl of chloroform (Fisher Scientific; Cat# C/4960) was added to achieve phase separation into three different layers. The aqueous upper phase, containing total RNA was transferred to a sterile tube and the rest of the protocol was carried out according to manufacturer's instructions. In the final step, RNA pellet was resuspended in sterile, RNase-free water (Merck Millipore; Cat# H20MB0501) and samples were stored at -20 ºC.

A DNase treatment step was carried out to ensure there was no contamination of the sample with genomic DNA. For this purpose, 0.1X sample volume of 10X TURBO DNase Buffer (Ambion; Cat# AM2238) and 1 µl TURBO DNase 2 U/ µl (Ambion; Cat# AM2238) were added to each sample. The mixture was incubated at 37 ºC for 30 min. To stop the reaction, EDTA was added to a final concentration of 5 mM and samples were heated to 70 ºC for 5 min. RNA was then quantified using the NanoDrop 2000 UV-Vis (Thermo Scientific).

2.3.9. Reverse transcription

To make cDNA from total extracted RNA, the iScript cDNA synthesis kit (Biorad; Cat# 1708890) was used according to manufacturer's instructions. 1 µg total RNA that had undergone DNase treatment was used for each reaction. A thermal cycler (G-Storm 482) was used for the sequential incubations. Samples were either stored at 4 ºC overnight short term, or at -20 ºC.

2.3.10. *In vitro* transcription (IVT) and RNA purification

Depending on the application, different IVT kits and protocols were followed:

1) To generate digoxigenin (DIG)-labelled probes for *in situ* hybridisation, the DIG RNA Labelling Kit (SP6/T7) (Roche; Cat# 11175025910) was used. Per 20 µl reaction, set up in a sterile, RNase free tube, the following components were added in this order: 0.9-1 µg of linearised plasmid, RNase-free water to 13 µl, 2 µl NTP Labelling Mixture (10X concentrated), 2 µl Transcription Buffer (10X concentrated, at room temperature to avoid precipitation), 1 µL Protector RNase Inhibitor (20 U/µl) and 2 µl SP6 or T7 RNA polymerase (20 U/µl). The reaction was mixed gently but thoroughly and was incubated for 2-3 h in a 37 ºC incubator. 2 µl DNase I (10 U/µl) was added and the reaction was incubated for another 15 min at 37 ºC. To stop the reaction, 1 µl of 0.5 M EDTA, DEPC-treated water to 100 µl, 10 µl of 4 M lithium chloride (LiCl, RNase-free) and, finally, 2.5 volumes of cold 100% EtOH were added. RNA was precipitated at -20 ºC overnight.
2) To generate CRISPR gRNA, the MEGAshortscript™ T7 Transcription Kit (Ambion; Cat# AM1354) was used according to manufacturer's instructions. 8 µl of template DNA was added in each 20 µl reaction. To achieve maximum yield, reactions were incubated at 37 ºC for several hours or overnight. Typically, 5 h of incubation produced 36 µg of RNA following precipitation, whereas overnight incubation produced on average 450 µg. Following TURBO DNase treatment, reactions were processed with ammonium acetate stop solution (provided in the kit), according to manufacturer's instructions. RNA was precipitated at -20 ºC overnight.
3) To generate Cas9 mRNA and mRNA used for the overexpression studies (refer to section 2.2.3), the mMessage mMachine SP6 Transcription Kit (Ambion; Cat# AM1340) was used according to manufacturer's instructions. Up to 1 µg of linear plasmid template was added per 20 µl reaction. Reactions were incubated at 37 ºC for 2-3 h, followed by a TURBO DNase treatment step. The synthesised RNA was recovered using LiCl precipitation solution (provided in the kit). RNA was precipitated at -20 ºC overnight.
4) To generate the *tfec* construct for overexpression, the mMESSAGE mMACHINE T7 ULTRA Transcription Kit (Ambion; Cat# AM1345) was used according to manufacturer's instructions. The kit contained reagents that facilitate mRNA capping and drive *in vitro* assembly of a poly(A) tail. To generate mRNA from a PCR amplicon, 100 µg of purified DNA template were added to a 20 µl reaction, which was incubated for 2 h at 37 ºC in a thermal cycler (G-Storm 482). The poly(A) tailing reaction (set up according to manufacturer's instructions) was

incubated for 45 min at 37 ºC. The mRNA construct was purified using LiCl stop solution (provided in the kit).

Following overnight precipitation, samples were centrifuged for 30 min at 13,000 rpm at 4 ºC to pellet the RNA, which was then washed with 500 µl ice cold 70% EtOH. The samples were then centrifuged for 10 min at 13,000 rpm at 4 ºC. The EtOH was removed and the pellet was air dried for 3-5 min at room temperature, before being resuspended in molecular biology grade water (Merck Millipore; Cat# H20MB0501).

The NanoDrop 2000 UV-Vis Spectrophotometer (Thermo Scientific) was used to quantify and assess RNA purity based on the value of the 260/280 ratio (approximately 2.0 for pure RNA). The RNA was diluted and aliquoted as necessary for each application. A single aliquot was used to run on an agarose gel (refer to section 2.3.11) to confirm RNA integrity. Long-term storage of RNA aliquots was at -80 ºC.

2.3.11. Agarose gel electrophoresis for RNA

Following IVT, it is essential to assess the integrity of synthesised RNA. To that end, aliquots containing 100-250 ng RNA were, if necessary, diluted with DEPC-treated water and an appropriate volume of 2X RNA loading dye (NEB; Cat# B0363S) was added. Samples were heated to 70 ºC in a thermal cycler for 5 min and were electrophoresed for approximately 30 min at 100 V on a 1% agarose gel made using TBE buffer. The tank in which the gel was run, as well as the gel mold and comb, were bleached prior to use. The approximate size of RNA fragments was calculated based on a ssRNA ladder (NEB; Cat# N0362S), which was processed according to manufacturer's instructions and run alongside the samples on the gel.

2.3.12. Solutions

<u>50 mM NaOH</u>

0.2 g NaOH pellet (Fisher; Cat# S318) were dissolved in 50 ml sterile water to generate 100 mM NaOH. This solution was then diluted 1:2 to generate 50 mM NaOH used for extraction.

The solution was prepared fresh prior to extraction.

<u>Proteinase K solution</u>	<u>1X TAE</u>	<u>1X TBE</u>
Final concentration in sterile water:	40 mM Tris (pH 7.6)	89 mM Tris (pH 7.6)
	20 mM acetic acid	89 mM boric acid
10 mM TrisHCl pH 8	1mM EDTA	2 mM EDTA
200 µg/ml Proteinase K	in sterile water.	in sterile water

2.4. Cloning techniques

2.4.1. Ligation

For ligation of blunt-ended PCR products into plasmid vectors, the Zero Blunt TOPO Vector System (Invitrogen; Cat# 450245) was used. 4 µl of freshly purified PCR product were combined with 1 µl salt solution and 1 µl of pCR II-Blunt-TOPO vector (for plasmid map refer to Appendix section I.4). The reaction was mixed gently and placed in a 23 ºC incubator for 15-20 min before placing on ice. 3 µl of the ligation were subsequently used for transforming DH5α competent cells (refer to section 2.4.2).

For PCR fragments with T-overhangs the pGEM T Easy Vector System (Promega; Cat# A1360) was used (for plasmid map refer to Appendix section I.3). The following reagents were combined: 5 µl of 2X Rapid Ligation Buffer, 1 µl of pGEM T Easy Vector (50 ng/µl), 3 µl of purified PCR product and 1 µl of T4 DNA Ligase (3 U/µl). The ligation was allowed to progress overnight at 4 ºC to achieve optimal yield. To transform DH5α competent cells, 5 µl of this reaction were used.

2.4.2. Transformation

Plasmid vectors were transformed into competent *E. coli* cells by heat shock. Glycerol stocks of the commercially available DH-5αF' strain (Clontech Laboratories Inc., CA) were used to prepare 100 µl transformation aliquots in the lab using chemically-induced competency methods. Competent cells were stored at -80 ºC. To transform, the appropriate number of 100 µl aliquots were defrosted on ice for 20 min. Either the appropriate ligation reaction volume (refer to section 2.4.1), or 1 µl of a purified plasmid stock were added and cells were incubated on ice for 30 min with occasional gentle tapping. The tubes were incubated in a water bath set to 42 ºC for 45 sec and were immediately placed on ice for 2 min. After transformation, cells were incubated in SOC medium for 1.5 h, prior to spreading on Luria broth (LB) agar plates containing a suitable antibiotic (see below).

2.4.3. Bacterial cultures

LB agar plates were prepared by mixing warm liquid agar with the required antibiotic. Either Carbenicillin/Ampicillin or Kanamycin plates were used for routine propagation of glycerol stocks or of transformed bacteria. Antibiotics were used at a final concentration of 50 µg/µl. For identification of colonies transformed with successfully ligated pGEMT vector, plates containing Xgal were used for blue-white screening. 1 mg of Xgal was spread on plates containing antibiotic, and was allowed to be absorbed for 10-20 min. IPTG was added directly to the transformation sample at a final concentration of 0.34 mM prior to plating. Bacteria were streaked (glycerol stocks) or spread (transformations)

on the agar plates with the appropriate sterility precautions. Plates were incubated at 37 ºC overnight.

Prospective positive colonies were picked and dissolved in 2-5 ml of LB medium with the appropriate antibiotic (final concentration 50 µg/µl). Cells were grown at 37 ºC overnight in a shaking incubator (started culture). To obtain bacteria for a midiprep, 1 ml of starter culture was added to 100 ml of freshly autoclaved LB medium containing the appropriate antibiotic.

2.4.4. Plasmid isolation

Plasmid isolation from bacterial cultures between 2-5 ml was done using the GeneJET Plasmid Miniprep Kit (Thermo Scientific; Cat# K0502) according to manufacturer's instructions. Plasmid midiprep preparations were carried out using the NucleoBond Xtra Midi kit (Macherey-Nagel; Cat# 740410) according to manufacturer's instructions. In both cases, plasmids were eluted in sterile water and stored at -20 ºC.

2.4.5. Solutions

LB medium

2.5 % (w/v) Luria broth base in Milli-Q water

LB agar

3.7 % (w/v) agar base in Milli-Q Water

2.5. *In situ* hybridisation

2.5.1. Probe preparation

Probe design and synthesis

The probe used for *tfec* transcript detection was designed *de novo*, using the Ensembl database to obtain the cDNA sequence of the gene and the Primer3 tool to design primers that would amplify a 946 bp region of the ORF. The specificity of the primers against the zebrafish genome was confirmed using Primer Blast. The remainder of the probes have previously been designed either by members of our group, or by other research groups. For more details regarding probe sequences and respective vectors refer to table 2.2.

To generate probes for *tfec* and *ltk* transcript detection (primers for the latter were designed by Susana Lopes, PhD thesis), the following procedure was followed: Total RNA was extracted from zebrafish embryos (refer to section 2.3.8) at a stage when the genes of interest are prominently expressed. Total RNA was used as a template to

synthesise cDNA using the SuperScript III First-Strand Synthesis System (Invitrogen; Cat# 18080051) according to manufacturer's instructions.

To amplify the fragment of interest, KOD Hot Start High Fidelity DNA Polymerase (Novagen, Cat# 71086) was used. Reactions were assembled according to manufacturer's instructions, adding 3 µl of cDNA template per reaction. Cycling conditions were as follows: initial denaturation at 95 ºC for 2 min, followed by 35-40 cycles of incubation for 20 sec at 95 ºC, 30 s at 55 ºC (annealing temperature varied according to primer set) and 1 min at 72 ºC, cycles were followed by final extension for 10 min at 72 ºC. Products were assessed by agarose gel electrophoresis (refer to section 2.3.3) to confirm that the expected size of amplicon was present.

Amplicons were purified by gel extraction (refer to section 2.3.5), ligated into Zero Blunt TOPO vector (Invitrogen; Cat# 450245) and transformed into DH5α competent cells (refer to sections 2.4.1 and 2.4.2). The transformations were plated on LB agar/kanamycin plates, colonies were picked the following day and inoculated overnight in LB/Kanamycin broth (refer to section 2.4.3). To purify plasmid DNA, the GeneJET Plasmid Miniprep Kit (Thermo Scientific; Cat# K0502), was used according to manufacturer's instructions. The concentration of DNA was quantified using the NanoDrop and samples were sent for sequencing (refer to section 2.3.6). Standard nucleotide BLAST (nBLAST) was used to confirm the identity of the sequence cloned in the vector, as well as to determine the orientation of the fragment. To synthesise RNA complementary to the endogenous mRNA of the gene of interest (antisense probe), IVT was initiated from the opposite end of the sequence compared to *in vivo* transcription.

For all probes used, 5 µg of plasmid was linearised by restriction digestion adjacent to the 3' end of the sequence used as a probe. Restriction digests were allowed to progress for 2- 4 h depending on the nature of each enzyme. Agarose gel electrophoresis was used to confirm that no undigested plasmid was present in the sample. Phenol-chloroform-isoamyl alcohol extraction (refer to section 2.3.5) was then used to purify the linearised plasmid. The DIG-labelled RNA probes were synthesised and purified as described in section 2.3.10 (1), and construct integrity was validated by agarose gel electrophoresis.

Probe hydrolysis

To increase permeability, probes for *tfec*, *ltk* and *pnp4a* transcript detection were fractionated to a final size of approximately 600 nucleotides. For this, 50 µl of intact probe in DEPC-treated water, 30 µl DEPC-treated water, 10 µl of 0.4 M sodium bicarbonate and 10 µl of 0.6 M sodium carbonate were mixed in a sterile RNase-free microfuge tube.

The solution was heated in a water bath at 60 ºC for times calculated using the following equation:

$$time\ (min) = \frac{(\text{starting kb} - \text{desired kb})}{(0.11 \times \text{starting kb} \times \text{desired kb})}$$

Hydrolysed probes were precipitated by adding 50 µl DEPC-treated water, 10 µl of 3 M sodium acetate pH 4.5, 1.3 µl glacial acetic and 300 µl 100% EtOH. The mixture was placed at -20 ºC overnight and precipitated the following day as described in section 2.3.10. To perform *in situ* hybridisation, 1:50 and 1:100 dilutions of hydrolysed probe stocks in hybridisation mix (HM) were maintained at -20 ºC.

Pre-adsorption

Prior to performing *in situ* hybridisation, probes were pre-adsorbed to increase the signal: background ratio by eliminating non-specific binding. For this, the first day of the *in situ* hybridisation protocol (refer to section 2.5.3) was followed for samples of approximately 50 embryos per tube. Embryos were not bleached. The hybridisation step was extended for approximately 24 h and then the probe was recycled, while the embryos were discarded.

Table 2.2. Detailed information regarding the generation of *in situ* hybridisation probes.

Gene name	Plasmid backbone	Reference	Sequence length	Linearisation enzyme	IVT enzyme
tfec	Zero-Blunt TOPO	This study	946 bp	*PstI*	SP6
ltk	Zero-Blunt TOPO	(Lopes et al. 2008)	1,3 kb	*NotI*	SP6
mitfa	pBluescript	(J A Lister et al. 1999)	1.6 kb	EcoRI	T7
foxd3	likely pBluescript SK+	(Odenthal and Nüsslein-Volhard 1998)	970 bp	*BamHI*	T7
pnp4a	pCR-BluntII TOPO+	(Curran et al. 2010)	1.4 kb	*EcoRV*	SP6
sox10	pGEMT	(Dutton et al. 2001)	2 kb	*SalI*	T7

2.5.2. Sample preparation

Approximately 25-30 zebrafish embryos were transferred to a 2 ml microfuge tube. The embryo medium was removed and samples were initially washed with approximately 400 µl of 4% PFA. The fixative was removed and 1 ml of 4% PFA was added to each tube. Samples were fixed either at room temperature for 4 h, or at 4 ºC overnight, in both cases while being gently agitated. Fixed embryos were first washed for 5 min with 1 ml PBT, then with 1 ml with each of 25% MeOH/75% PBT, 50% MeOH/50% PBT, 75% MeOH/25% PBT and 100% MeOH solutions. Each wash was for 5 min at room temperature, with slow agitation. Ultimately, fresh 100% MeOH was added and samples were stored at -20 ºC at least overnight or as long as required before use.

Embryos <22 hpf were not dechorionated prior to fixation. Instead, following fixation, they were washed twice with PBT and transferred to a Petri dish filled with PBT using a plastic Pasteur pipette. They were then dechorionated before dehydration took place by subsequent MeOH washes. Embryos >22 hpf were dechorionated and allowed approximately 30 min for the trunks to straighten prior to being collected for fixation.

2.5.3. *In situ* hybridisation protocol

All washes of the *in situ* hybridisation protocol took place at room temperature with slow agitation, using 1 ml of solution, unless stated otherwise. Washes at 67.5 ºC were carried out in a water bath. For full descriptions of solution composition refer to section 2.5.5.

Day 1

Samples were rehydrated for 5 min in each of the following solutions: 75% MeOH/25% PBT, 50% MeOH/50% PBT, 25% MeOH/75% PBT and twice in PBT. For embryos >24 hpf that had not been PTU treated, 600 µl of bleaching solution was added per sample and, depending on the stage and the degree of melanisation, the embryos were allowed 3-7 min at room temperature, with the tubes positioned horizontally on the bench. To fully remove bleaching solution, the embryos were rinsed once and then washed twice for 5 min in PBT. Embryos that had not undergone bleaching were washed twice for 5 min in PBT.

To permeabilise the embryos, proteinase K (10 mg/ml) was diluted in PBT to a final concentration of 1 µg/ml. The solution was applied and samples were incubated at 37 ºC without agitation according to the following table:

Table 2.3. Permeabilisation times for *in situ* hybridisation.

Stage	18 hpf	24 hpf	30 hpf	36 hpf	48 hpf	60 hpf	72 hpf
Permeab. time	2 min	5 min	6 min	8 min	9 min	12 min	15 min

Samples were rinsed with PBT and re-fixed in 4% PFA at room temperature with agitation for 20 min to 1 h. They were then washed 4X 5 min in PBT before hybridisation mix (HM) was applied. Pre-hybridisation in HM took place at 67.5 ºC for 5-7 h. The HM was removed and 70-200 µl of probe solution were added to each sample. Hybridisation was carried out overnight at 67.5 ºC (or at 70 ºC for the *ltk* probe). To generate probe solutions, stocks were diluted 1:100 in HM, with the exception of the *ltk* probe solution, which was made by diluting the stock 1:500.

Day 2

All solutions used for washes at 67.5 ºC were pre-warmed for 10 min in the water bath. To remove excess and non-specifically bound probe, the samples were submitted to the following washes at 67.5 ºC: rinse with 500 µl of 100% HM, followed by 10 min wash with each of 75% HM/25% 2X SSCT, 50% HM/50% 2X SSCT, 25% HM/75% 2X SSCT and 2X SSCT. Then samples were washed twice for 30 min using 0.2X SSCT. The following washes took place at room temperature with gentle agitation: 5 min in each of 75% 0.2X SSCT/25% MABT, 50% 0.2X SSCT/50% MABT, 25% 0.2X SSCT/75% MABT and twice in 1X MABT. 1% blocking solution (BS) was applied and samples were washed with gentle agitation at room temperature for 5-7 h. Finally, anti-DIG alkaline phosphatase-binding antibody (Roche, Cat# 11093274910) diluted 1:5,000 in 1% BS was added and samples were incubated at 4 ºC overnight with slow agitation.

Day 3 onwards

Samples were rinsed with 1X MABT. Then a series of 3X 5 min and 3X 10 min washes in 1X MABT took place, followed by 3X 5 min washes in colouration buffer (CB). Embryos were transferred in glass 9 well plates, using a Pasteur pipette, and incubated in NBT/BCIP solution for 30min to 3 days, until the signal: noise ratio became optimal. To stop the reaction, embryos were rinsed in PBT, transferred to clean microfuge tubes and washed in PBT for 5 min. If a significant level of background was present, the embryos were immersed in 100% MeOH and incubated at 37 ºC for 10 min. This step achieves better contrast for imaging purposes but was not conducted in samples where background was not present. If MeOH treatment took place, samples were washed twice for 5 min with PBT at room temperature. All samples were re-fixed in 4% PFA for 20-30 min at room temperature, with gentle agitation. For analysis, imaging and long term storage, samples were washed for 10 min in 30% glycerol, 10 min in 50% glycerol and 10 min in 80% glycerol (diluted in PBT) and were stored at -20 ºC at least overnight.

2.5.4. Analysis and microscopy

To analyse gene expression patterns, embryos in 80% glycerol were transferred into glass 9-well plates using a Pasteur pipette and were aligned using fine mounting needles and a pair of forceps. Initial phenotypic analyses and cell counts were performed under a standard dissecting microscope (Leica MZ75) using incident light. If a higher magnification was required for cell counting, the Zeiss Axio Zoom.V16 fluorescence stereo zoom microscope was used.

For imaging, embryos were mounted on glass slides (VWR; 6310114), between a pair of 18x18 mm coverslips glued approximately 1.5 cm apart. Single embryos were transferred in the space between the side coverslips using a Pasteur pipette, were oriented appropriately with a fine mounting needle and a coverslip (18x18 mm or 22x22 mm depending on the stage of the mounted embryo) was gently placed on top. The position of the top coverslip was manually adjusted until the embryo was oriented laterally. If dorsal imaging was required, two different techniques were used: to achieve better optics at high magnification (20X, 63X objectives), the region of the embryo to be imaged was isolated and mounted as described above.

If lower magnification was sufficient, the embryo was mounted in 1% low melting point agarose (Invitrogen; Cat# 16520050) dissolved in sterile water. The embryo of interest was transferred on a glass slide, covered with a drop of agarose solution and manipulated with a pair of fine mounting needles until the agarose started to set. To provide further support, if necessary, using a plastic Pasteur pipette, a ring of agarose was drawn around the mounted embryo, the height of which should be the same as that of the agarose layer in which the embryo was embedded. After the agarose forming the ring had solidified, a longer coverslip was placed on top and, using a glass Pasteur pipette, sufficient agarose was pipetted from the sides in order to fill the gaps.

For imaging, the upright compound Zeiss Imager 2 microscope, fitted with an Axiocam 506 colour camera (Zeiss), was used. Images presented in this work were taken using the 10X, 20X and 63X (water dipping) objectives for brightfield and DIC images. The ZEN2 software (Zeiss) was used for acquiring and processing images, including Z-stacks.

2.5.5. Solutions

PBT

0.1% (v/v) Tween20 in PBS

PBS was prepared by dissolving commercially supplied tablets in Milli-Q water.

Bleaching solution

Dissolved in sterile water 1% (v/v) KOH 1.5% (v/v) H_2O_2

Hybridisation mix

Dissolved in DEPC-treated water: 50% (v/v) Formamide 5X SSC 50 ng/mL Heparin 500 ng/mL tRNA 0.1% (v/v) Tween20 9.2 mM Citric Acid

20X SSC

Dissolved in sterile water: 3M NaCl 300 mM sodium citrate pH was adjusted to 7.0 and the solution was autoclaved.

2X SSCT

1: 10 dilution of 20X SSC, addition of 0.1% (v/v) Tween20.

10X MAB

Dissolved in sterile water: 1 M Maleic acid 900 mM NaCl pH was adjusted to 7.5 and the solution was autoclaved.

1X MABT

1: 10 dilution of 10X MAB, addition of 0.1% (v/v) Tween20.

1% Blocking solution

Dissolved in 1X MABT:2 mg/ml BSA 5% (v/v) Normal Goat Serum

PFA

4% (w/v) paraformaldehyde in PBS

DEPC-treated PBS was warmed to 65 ºC in a water bath. PFA powder was added and the pH was adjusted to 7.0-7.5. Once the PFA was fully dissolved, 50 ml aliquots were prepared and stored at -20 ºC.

Colouration buffer

Dissolved in sterile water: 100 mM Tris HCl pH 9.5 100 mM NaCl 0.1% (v/v) Tween20

NBT/BCIP solution

Dissolved 5 ml NBT/BCIP stock solution (Roche; Cat# 11681451001) in 100 ml colouration buffer.

2.6. RNAscope multiplex fluorescent assay

RNAscope assays allow for fluorescently labelling a maximum of three gene transcripts per sample. The technique has been relatively recently applied using whole-mount zebrafish embryo samples (Gross-Thebing et al. 2014). Figure 2.2 (adapted from Gross-Thebing et al.) diagrammatically represents the molecular steps involved in fluorescently labelling distinct mRNA molecules in the same cell. RNAscope presents the advantages of increased sensitive, compared to conventional transcript detection approaches, such as *in situ* hybridisation, and of allowing for relative quantitation of transcripts per cell. Using the appropriate imaging equipment and image analysis software, the intensity and extent of fluorescent staining in different cells can be used to indicate number of transcript molecules in those cells.

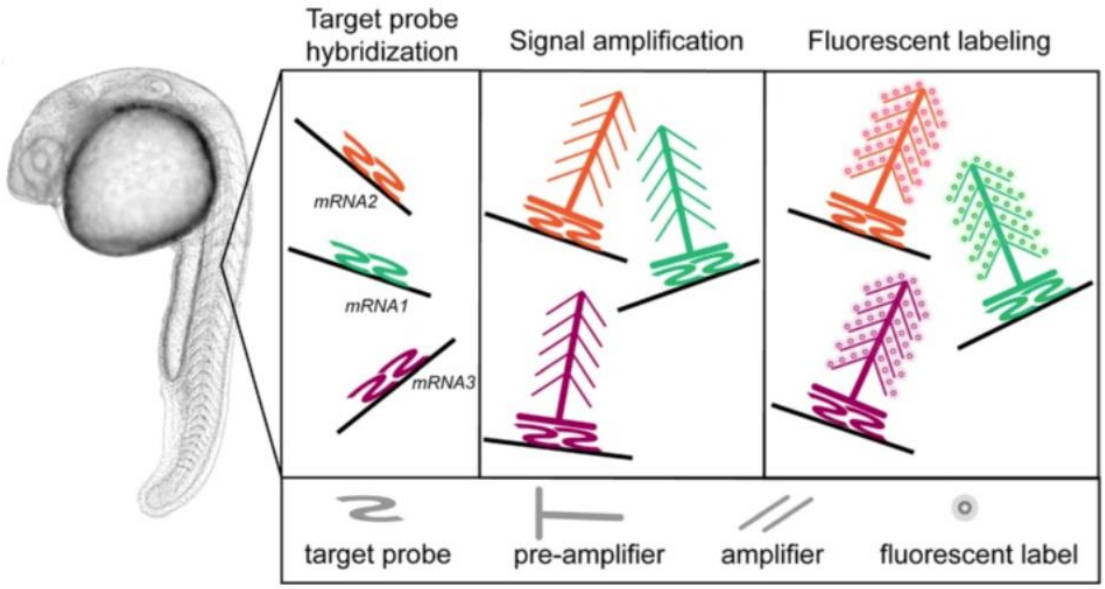


Figure 2.2. The principle of RNAscope. First, commercially designed and synthesised z-shaped oligonucleotide probes specifically bind adjacent to each other to their respective transcript sequences, forming ‘zz’ pairs. Second, commercially manufactured pre-amplifier and amplifier molecules are added and ‘zz’ pair-specific scaffolds assemble on the surface formed by each pair. Finally, fluorescent compounds are added, recognise amplifier molecules and emit fluorescence in different channels. Figure adapted from (Gross-Thebing et al. 2014).

2.6.1. Sample preparation

To obtain samples for RNAscope, approximately 12 WT embryos, dechorionated as described in section 2.5.2, were placed in 2 ml microfuge tubes, to ensure homogeneous treatment. The embryo medium was discarded and replaced by 1 ml of 4% PFA (refer to section 2.5.5). Fixation took place at room temperature for 1 h with slow agitation.

Then embryos were washed three times for 5 min with 1 ml PBT and were dehydrated via a series of washes with 25%, 50% and 75% MeOH diluted in PBT, and then with 100% MeOH (5 min each wash). Fresh MeOH was added and the samples were placed at -20 ºC at least overnight, or for long-term storage.

2.6.2. Preparation of materials

All RNAscope materials were stored at 4 ºC, according to manufacturer's instructions. The following reagents were prepared ahead of time:

1) 0.2X SSCT as the main wash buffer (refer to section 2.5.5)
2) 1X PBT (refer to section 2.5.5)
3) Mixtures of probes to detect different transcripts of interest. Commercially synthesised and supplied probe stocks (ACD, USA) were prewarmed for 10 min at 40 ºC in a water bath to dissolve precipitation, before being brought to room temperature. Desired mixes of probes were prepared according to manufacturer's instructions. Mixtures of probes were stored at 4 ºC and reused.

2.6.3. RNAscope protocol

All washes of the RNAscope protocol took place at room temperature with very slow agitation, using 1 ml of solution, unless stated otherwise. Washes at 40 ºC were carried out in a water bath without agitation.

Day 1

Embryos were rehydrated via a series of 5 min washes in each of the following solutions: 75% MeOH/25% PBT, 50% MeOH/50% PBT, 25% MeOH/75% PBT and three times in PBT. To permeabilise the embryos, 2 drops of Pretreat 3 ready-to-use solution (ACD, USA; Cat# 320842) were added per tube and the samples were slowly agitated either for 15 min (for 18-24 hpf embryos) or for 20 min (for older embryos). The solution was removed and the embryos were washed three times for 5 min with PBT. A mixture of probes was prewarmed at 40 ºC for 10 min and was briefly allowed to cool at room temperature. PBT was fully removed from the samples and approximately 100 µl of probe mixture were added. Samples were incubated at 40 ºC overnight.

Day 2

The probes were recovered in clean tubes and stored at 4 ºC. Embryos were washed once for 5 min with 2X SSCT, then three times for 15 min with 0.2X SSCT. They were then re-fixed with 4% PFA for 10 min while agitating and washed three times for 15 min with 0.2X SSCT. Upon initiation of the second wash, the preamplifier hybridisation ready-to-use solution (Amp1; ACD, USA; Cat# 320851) was transferred from 4 ºC to room

temperature, to allow any precipitation to dissolve. After the final wash was finished, the SSCT was removed and 2 drops of Amp1 (Amp solutions were supplied in squeeze bottles) were added per sample. The tubes were lightly tapped to ensure homogeneous treatment of the embryos and placed at 40 ºC for 30 min.

Embryos were washed three times for 15 min with 0.2X SSCT. Following initiation of the second wash, the signal enhancement ready-to-use solution (Amp2; ACD, USA; Cat# 320851) was brought to room temperature. After the final wash was finished, the SSCT was removed and 2 drops of Amp2 were added per sample. The tubes were gently tapped and placed at 40 ºC for 15 min.

Embryos were again washed three times for 15 min with 0.2X SSCT. In the duration of the washes, the amplifier hybridisation ready-to-use solution (Amp3; ACD, USA; Cat# 320851) was brought to room temperature. The SSCT was removed and 2 drops of Amp3 were added. The tubes were lightly tapped and placed at 40 ºC for 30 min.

The series of three 15 min washes with 0.2X SSCT was repeated. During that time, the appropriate ready-to-use solution containing the fluorescent compounds (Amp4 AltA/B/C; ACD, USA; Cat# 320851) was brought to room temperature. Alternative (Alt) versions of Amp4 allow for probes of each channel to be represented by different fluorophores, thus expression of each gene could be later visualised with a convenient optical filter for each application. After the final wash, the SSCT was removed and 2 drops of the appropriate Amp4 solution were added. The tubes were gently tapped and placed at 40 ºC for 30 min. At this stage, exposure of samples to light was minimised to prevent photobleaching of the fluorophores.

Samples were washed three times for 15 min with 0.2X SSCT. After the second wash was initiated, the commercially supplied ready-to-use DAPI solution (ACD, USA; Cat# 320851) was brought to room temperature. The wash buffer was removed and 2 drops of DAPI were added to each sample. The tubes were lightly tapped and placed at 4 ºC overnight with slow agitation.

2.6.4. Microscopy and analysis

Embryos to be imaged were first transferred to a glass well containing 80% glycerol. Exposure to light was minimised to prevent photobleaching. Single embryos were mounted for imaging by transferring them on a glass slide (VWR; 6310114) in 80% glycerol. A single side coverslip was used for 18 hpf sample, to prevent disintegration of the embryo. For older samples, no side coverslips were used because the strength of DAPI fluorescence resulted in increased imaging artefacts when embryos were not compressed. An added advantage was that Z-stacks could capture a greater depth of field before photobleaching set in.

Images were taken with the upright compound Zeiss Imager 2 microscope, using the monochrome Orca Flash 4.0 V2 camera (Zeiss) and Apotome.2 (Zeiss) for optical sectioning. For fluorescent images, dsRed (for the Alexa 550 dye), YFP (for the Alexa 488 dye) and DAPI filters were used (supplied by Zeiss) at 20X magnification. 3 optical sections were taken per Z-slice. To decrease imaging time, maximum light intensity and minimum possible exposure time were used.

Images were captured and analysed with the ZEN software (Zeiss), before being processed with ImageJ (FIJI package). For each developmental stage, the same region of the anterior trunk was imaged.

It has been previously reported that in samples subjected to RNAscope, individual fluorescent spots observable in the cytoplasmic compartment represent single mRNA molecules (F. Wang et al. 2012). Using the Z-stack images obtained as described in this section, relative quantitation of gene expression in single cells of whole mount zebrafish embryos was achieved by scoring fluorescent spots located in close proximity to the periphery of individual DAPI-stained nuclei. It should be noted that only relative quantitation was possible, as the achievable resolution was deemed likely inadequate to resolve spots corresponding to single transcripts. It was thus considered possible that, especially in cells presenting with high gene expression levels, several single spots positioned close to each other were visualised as single unresolved spots.

Gene co-expression was determined using the Z-stack files by (1) identify nuclei of cells surrounded by signal which corresponds to one of the genes of interest, (2) assessing the area surrounding each individual nucleus for fluorescent spots indicating expression of the second gene. Although visible overlap of fluorescent signals is strongly indicative of co-expression, in cells presenting with low expression levels, this would be less likely. Thus, in these cases co-expression was evaluated based on proximity of the different-coloured spots to a single nucleus. Where the tissue morphology and image quality allowed it, DIC Z-stacks were used to define cell boundaries, in order to ensure both genes investigated were indeed expressed in the same cell, rather than in adjacent cells.

A challenge when analysing RNAscope results was to distinguish specific signal, indicating gene expression, from non-specific background and autofluorescence. Non-specific signal commonly arose from trapping of the probes within the notochord, resulting in strong fluorescence in that region. To avoid misinterpretation of gene expression due to this phenomenon, cells proximal to the notochord, such as medially migrating NC derivatives, were excluded from analyses. For the same reason, cells located very close to tissues that presented with high levels of gene expression for a gene of interest were not assessed for expression of that specific gene. For example, *foxd3* expression was not evaluated in medially migrating cells due to high levels of *foxd3*

expression in overlying somites. Moreover, in some cases non-specific binding of the probes can be distinguished as the resulting fluorescent spots do not clearly surround distinct nuclei, but are instead evenly distributed along the embryo or along a specific structure, for instance somite boundaries. Finally, where high exposure times were required for imaging, arising autofluorescence could interfere with signal recognition. Specific signal could be distinguished from tissue autofluorescence owing to its punctuate appearance as opposed to the homogeneous nature of autofluorescence, which, additionally, was identified as it is detectable in more than one channels.

2.7. Mathematical modelling

2.7.1. Derivation of differential equations

The relationships between genes in the GRNs described were mathematically modelled using a system of ordinary differential equations (ODEs), which was derived according to instructions previously published by Greenhill et al. (Greenhill et al. 2011).

We defined the following terms in a population of identical cells:

$[G^P]$: The average concentration of protein G^P, encoded by gene G.

$[r_A^G]$: The average number of elements in the promoter of gene G that have been bound by A^P (normalised to the total number of gene G promoter elements recognisable by A^P).

$[G_i^*]$: The average number of transcriptionally active genes per cell, normalised to the total number of active and inactive genes.

Moreover, we defined the following rates:

a_A^G*:* The rate of binding of A to the promoter of G.

b_A^G*:* The rate of unbinding of A from the promoter of G.

Kd_A^G*:* The dissociation constant at equilibrium. By definition $Kd_A^G = b_A^G/a_A^G$.

g_G*:* The maximum rate of gene G expression. This rate incorporates mRNA and protein production levels, which were not explicitly modelled in this network.

d_G*:* The degradation rate of G^P.

The rate of change of *sox10* input was manually set using a Heaviside function (http://uk.mathworks.com/help/symbolic/heaviside.html). Physiologically relevant parameter values were identified from the literature and initial values were deduced based on detectable levels of gene expression via *in situ* hybridisation in the embryonic posterior dorsal trunk (refer to section 3.3.3).

2.7.2. Computational simulation

The software MATLAB was used to numerically solve the system of ODEs. ode45 was used as a solver, relative and absolute tolerance values were set to 10^{-4}.

2.8. Statistical analysis

2.8.1. Pearson's chi-squared tests

Pearson's chi-squared (χ^2) tests (Pearson 1900; Harris 1912; Griffiths et al. 2000) were used to confirm that phenotypic differences identified amongst *in situ* hybridisation samples were not due to chance or experimental artefacts, but were caused by loss of gene function. Therefore, our null hypothesis stated that the number of embryos displaying each of the observed phenotypes reflected expected Mendelian ratios (1 mutant: 3 WT embryos for recessive alleles).

The χ^2 values were calculated using the following formula:

$$\chi^2 = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

For these experiments, degrees of freedom = 1. We used the chi-square table (Jones 2008) to calculate the probability of deviations between expected and observed ratios being due to artefacts. We set our threshold probability to 0.01, meaning that only if the probability calculated from the table is larger than 0.01, deviations are due to chance. Pearson's chi-square test is deemed reliable for our samples because all of them contain over 10 embryos. It should be noted that the test only applies to numbers of progeny, not percentages.

2.8.2. Two sample t-test

Two sample t-tests were used to determine whether the mean values of two independent groups significantly differed. Such tests were performed using the software Minitab. Means were considered statistically different if the calculated p-value was less than 0.05.

2.8.3. Linear regression

To measure the strength of association between the expression levels of two genes, as determined based on RNAscope experiments, we used simple linear regression analysis performed using Minitab software. The strength of association was primarily determined using the r^2 value. Another measure of association is provided by the S value (standard error of the regression), which represents the average distance between the observed values and the predicted regression line. Finally, the calculated p-value indicates the statistical significance of the association.

3

The preliminary gene regulatory network

3.1. Introduction

This chapter aims to determine relationships between the literature-derived candidate genes of interest for the iridophore GRN. To this end, whole-mount chromogenic *in situ* hybridisation was primarily employed to spatiotemporally characterise the expression patterns of each gene in WT embryos and in siblings presenting with loss of gene function. These experiments allowed for discriminating between cell populations presenting with gene expression and for analysing loss of function effects specific to populations of interest. Based on the identified gene regulatory interactions, a preliminary GRN guiding iridophore specification from the NC (model A) was generated and mathematically modelled using a system of ordinary differential equations (ODEs). Using this mathematical expression of the network, predictions regarding the gene expression dynamics in model A were computationally produced.

To identify candidate members of the iridophore GRN, a thorough literature search was first performed. The candidate genes of interest were identified based on their previously established expression patterns and mutant phenotypes in relation to the embryonic iridophore population. Specifically, it was hypothesised that genes functioning towards iridophore specification would be either transiently or constitutively expressed in cells of this lineage. Additionally, loss of their function through mutagenesis or gene knock-down was expected to result in ablation of iridophore precursors.

The TF *sox10* is strongly expressed in embryos from as early as the 3-somite stage in cNCCs (Rau, Fischer, and Neumann 2006) and maintains high expression levels in the premigratory NC domain (Dutton et al. 2001). Previously conducted loss of *sox10* function studies identified the gene as crucial for specification of all NCCs, including iridophores that are eliminated in relevant mutants (R N Kelsh et al. 1996; Lopes et al.

2008). Expression of *sox10* in the iridophore lineage has thus far not been reported (Dutton et al. 2001; Greenhill et al. 2011; Takada and Appel 2011).

The RTK Ltk has been shown to specify a subpopulation of NCCs towards the iridophore fate (Lopes et al. 2008), whereas its role in the tetrapotent progenitor, termed chromatoglioblast, has been indicated but remains less clear (Nikaido et al., in prep.). In contrast to *sox10*, *ltk* expression in present throughout iridophore development (Lopes et al. 2008). Ltk has remained an orphan receptor for an extensive period of time (Lemke 2015) and it is only recently that efforts from several groups have focused on identifying the endogenous ligands of Ltk. The favoured and widely supported candidates are the mammalian cytokines FAM150A and FAM150B (Zhang et al. 2014; Guan et al. 2015; Reshetnyak et al. 2015), expression of which has been detected in several organs, including the thyroid, the adrenal gland, the brain and components of the gastro-intestinal tract (Zhang et al. 2014). In zebrafish three proteins have been identified: FAM150ba, FAM150bb and FAM150A (Lemke 2015), however their ability to activate zebrafish Ltk has not been addressed, while their expression patterns and loss of function effects have yet to be characterised. Because the ligand identities and expression dynamics remain largely elusive, the preliminary GRN featured binding of Ltk to an unknown ligand (L) to trigger activation of gene expression.

The gene *tfec*, which belongs to the MiT gene family and encodes a basic bHLH-ZIP TF (Rehli et al. 1999), was another chosen candidate. Importantly, TFEC is evolutionarily related to and displays striking sequence homology with the master regulator of melanocyte development, MITF (Hallsson et al. 2004; Pogenberg et al. 2012). In zebrafish, *tfec* expression has been previously reported in the premigratory NC domain, but importantly also in cells with spatiotemporal distributions reminiscent of specified and differentiated iridophores throughout development (James A Lister et al. 2011). Moreover, *tfec* morphants and CRISPR/Cas9-generated mutants demonstrated a striking loss of iridophores (J. Lister, unpublished data). The aforementioned evidence led to formulating the hypothesis that *tfec* likely functions as the iridophore master regulator. To confirm such a role it is important to establish the gene's function as not only necessary but also sufficient for iridoblast formation. Although this question was not addressed in this study, presented results aimed to establish the position of *tfec* in the iridophore GRN, subsequently relating it to that of *mitfa* in the melanophore GRN.

Finally, *mitfa* was included in the network. The gene is expressed in developing NCCs and in melanophores undergoing specification and differentiation. Interestingly, it has been shown that *mitfa* mutants have an increased number of mature iridophores (J A Lister et al. 1999). Furthermore, it has been reported that *mitfa* functions in the melanoiridoblast fate switch process (Curran et al. 2010). Such evidence indicates that

mitfa may engage in cross-repression with members of the iridophore GRN, an interaction which this study aimed to investigate further and which is further discussed in chapter 4.

To date, known regulatory relationships between these five genes of interest remain limited. Previous studies have shown that *sox10* regulates *mitfa* through direct binding on the gene's promoter (Potterf et al. 2000; Elworthy et al. 2003). Furthermore, it was shown that the *ltk* expression pattern is affected in *sox10* mutants (Lopes et al. 2008), although no conclusions were drawn with regard to *sox10*-dependent regulation of *ltk*. Results presented in this chapter aimed to establish the core GRN guiding iridophore development via thoroughly investigating the interactions between *sox10, tfec, ltk, mitfa* and *pnp4a* using both loss and gain of function assays.

3.2. Establishment of intergenic relationships

3.2.1. Spatiotemporal determination of iridophore development

For all loss of function experiments, embryos fixed at the following developmental stages were used: 18 hours post-fertilisation (hpf), 24 hpf, 30 hpf, 36 hpf and either 48 hpf or 60 hpf. These stages were considered to be most informative for obtaining a full understanding of the lineage's development. At 18 hpf, the majority of NCC progenitors remain restricted to the premigratory domain and are considered largely unspecified, or at the early stages of specification. At 24 hpf, NCCs express pigment cell markers while migrating ventrally, signalling that specification of these lineages is taking place. By 48 hpf mature iridophores can be readily observed. As has been previously demonstrated and was confirmed in this study, *tfec* expression in WT embryos, as visualised by *in situ* hybridisation (Fig. 3.1), highlights consecutive stages of iridophore development (James A Lister et al. 2011), and was therefore used to identify the relevant progenitor populations.

At 18 hpf, *tfec* expression was detectable in the majority, if not in all of the unspecified progenitors occupying the premigratory NC domain of the trunk (tNC) and the tail (Fig. 3.1 A). At this stage, *tfec* was expressed in the vicinity of the retinal pigment epithelium (RPE), in cNCCs and started to emerge within the aorta and the intermediate cell mass (ICM). The latter two locations correspond to sites where embryonic haematopoiesis takes place. It was recently shown that *tfec* is expressed in caudal endothelial cells and functions to regulate important cytokines, thus controlling the numbers and behaviour of haematopoietic stem cells (HSCs) (Mahony et al. 2016). For the purposes of this study, expression analyses were confined to *tfec* expression in the eye region and in NC derivatives.

At 24 hpf, a reduced number of cNCCs and no observable tNCCs expressed *tfec* (Fig. 3.1 B). However, expression remained strongly detectable in premigratory NCCs of the anterior tail, which is less developmentally progressed due to the nature of embryonic development, which occurs sequentially from anterior to posterior. In the anterior and middle trunk, *tfec* expression strongly labelled dorsally positioned and medially migrating cells (Fig. 3.1 B; Fig. 3.2 B), while it was also weakly detected in cells migrating via the lateral pathway (Fig. 3.2 A). Strong expression, referring to a cell or group of cells with prominent dark blue staining, was assumed to correspond to specified iridoblasts, whereas weak expression, identified by light blue colouration of a cell or group of cells, likely indicated NCCs becoming specified towards other lineages, such as the melanophore lineage, and hence downregulating *tfec*. Additionally, *tfec* expression at 24 hpf was detectable in the eye and in LP progenitors.

At 30 hpf, *tfec* expression formed a spotted pattern along the dorsal trunk of the embryo, with only the posterior region of the tail still showing the pattern corresponding to premigratory NCCs (Fig. 1C). Although specified iridoblasts occupy the same area of the dorsal trunk as the premigratory NCCs at an earlier stage, these different cell populations were readily distinguishable on the basis of their distribution. Specifically, NCCs are closely apposed and bilaterally distributed along the dorsal trunk, whereas iridoblasts form a series of isolated cells. At this stage, *tfec* remained strongly expressed in the eyes (data not shown) and in LP progenitors, as well as in medially migrating cells that were either undergoing specification, or had already become specified to form iridophores.

By 36 hpf, *tfec* expression was still detectable in persisting NCCs in the posterior-most region of the tail, the eyes (data not shown) and LP progenitors. Based on comparison with the established expression pattern of the iridoblast marker gene *ltk* (Lopes et al. 2008), *tfec* positive cells likely corresponding to specified iridoblasts were also identified along the dorsal and ventral trunk and the migrating pathways (Fig. 3.1 D). Finally, from 48 hpf onwards, *tfec* expression (Fig. 3.1 E,F) closely resembled differentiated iridophore positions (Fig. 1.2). Specifically, in keeping with previous reports (James A Lister et al. 2011), *tfec* expression was detected overlaying the RPE, LPs and the dorsal, ventral and yolk sac stripes, all of which are occupied by mature iridophores. Interestingly, *tfec* was also expressed in the epithelium of the swim bladder, which is not known to host iridophores (Fig. 3.1 F).

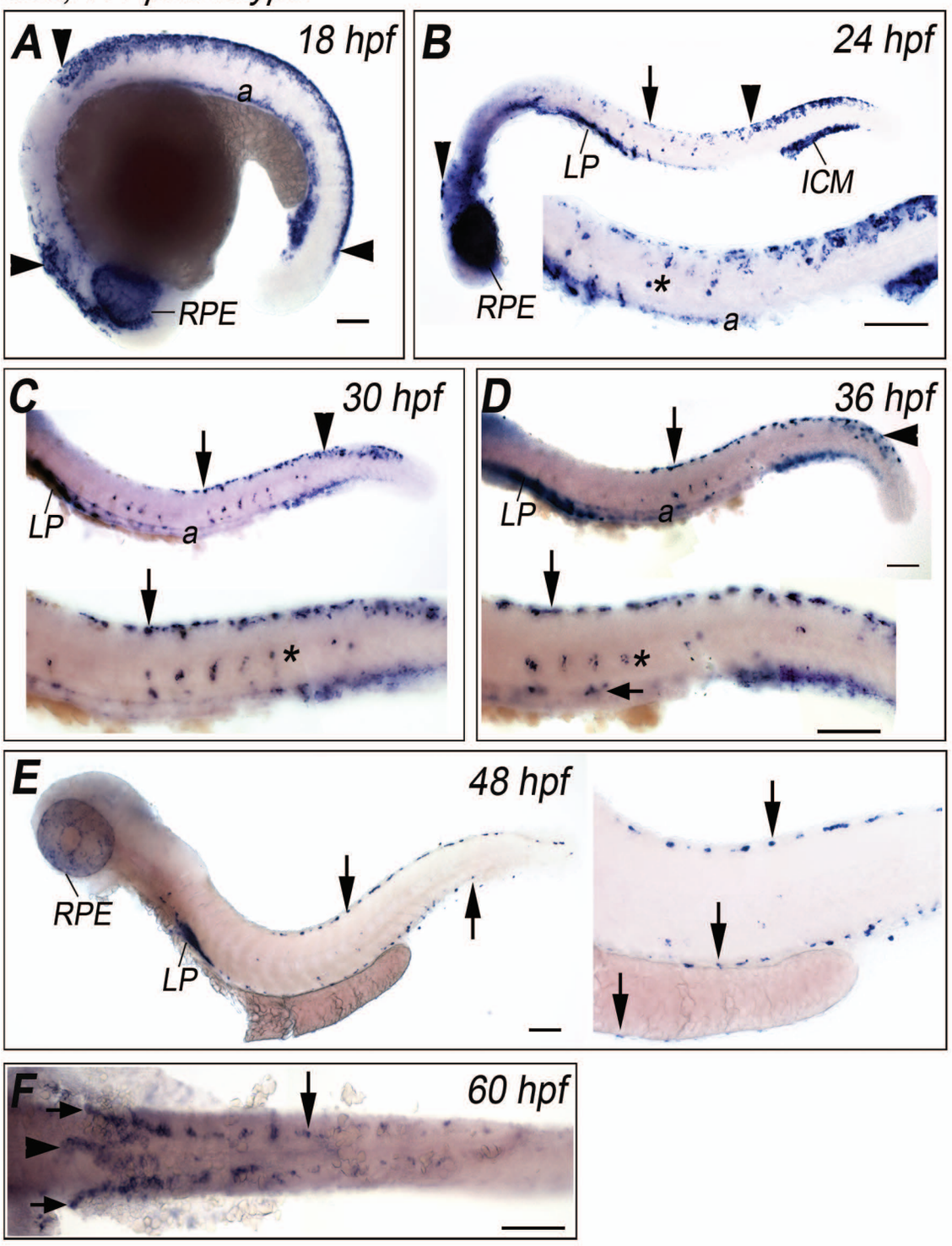


Figure 3.1. tfec expression labels cell populations relevant for iridophore development. *In situ hybridisation results show tfec expression (A) at 18 hpf in dorsally located cNCCs and tNCCs (arrowheads), LP progenitors, the RPE region, the aorta and the ICM. (B) At 24 hpf, in tNCCs (arrowhead), dorsally located specifying iridoblasts (arrow) and ventrally migrating NC derivatives (asterisk, insert). A reduced number of cNCCs still express the gene (arrowhead). (C, D) At 30 and 36 hpf, tfec expression is seen in dorsally and ventrally positioned (arrows) and migrating (asterisks) specified iridoblasts, very few cells of the premigratory NC domain (arrowhead) and LP progenitors. (E) At 48 hpf, tfec is expressed in mature iridophores of the dorsal, ventral and yolk sac stripes (arrows), developing LPs and overlaying the RPE. (F) At 60 hpf (dorsal view), the LPs (horizontal arrows), swim bladder epithelium (arrowhead) and ventrally positions iridophores (vertical arrow). RPE, retinal pigment epithelium; LP, lateral patches; a, aorta; ICM, intermediate cell mass. Lateral views unless stated otherwise, oriented with the head to the left. Scale bars: 100 µm.*

3.2.2. *sox10* regulates *tfec* in specified iridophores

To establish intergenic relationships between candidate genes of interest, whole-mount chromogenic *in situ* hybridisation experiments were performed using samples which comprised WT embryos, heterozygous mutant allele carriers and homozygous mutant siblings. Since embryos were derived from incrosses between identified heterozygous adults for each mutant allele, the three genotypes were expected to conform to Mendelian ratios (1:2:1). Unless stated otherwise, results from *in situ* hybridisation experiments in this study were interpreted by deriving p-values using the Pearson's χ^2 test, with the null hypothesis that fully recessive mutant phenotypes were observed (refer to section 2.8.1). These values indicated the likelihood that the number of embryos presenting with an alternative expression phenotype corresponded to approximately 25% of progeny and, therefore, that this phenotype likely resulted from loss of gene function. If the p-values were below 0.1, the phenotype was considered to not conform to expected ratios but to either arise from experimental artefacts or from non-recessive effects. Unless stated otherwise in figure legends, all figures indicate presumed WT and mutant phenotypes based on the p-values associated with observed ratios. Where presumed mutant phenotypes were mild and difficult to distinguish, embryos were molecularly genotyped.

To begin building the GRN governing iridophore specification, it was important to understand how *sox10*, a factor crucial for iridophore development, regulates iridogenic genes. First, *tfec* expression was assessed in *sox10*^{t3} mutants, compared to their WT siblings using *in situ* hybridisation. The iridophore lineage is largely eliminated in homozygous *sox10*^{t3} mutants, with very few differentiated cells remaining, known as 'escapers'. At 18 hpf, 100% of analysed embryos displayed the WT *tfec* expression pattern (N = 41; 2 independent experimental repeats; p-value < 0.001) (data not shown). From 24 hpf, expansion of the *tfec* positive NCC domain and absence of cells in the migratory pathways was observed in 20/74 embryos over 3 independent experiments (embryos with the alternative phenotype mutants according to Mendelian ratios; 0.6 < p-value < 0.7) (Fig. 3.3 A,B). At 30 hpf, the phenotype persisted in 13/48 of assessed embryos over 2 independent repeats (homozygous mutants according to Mendelian ratios; 0.7 < p-value < 0.8). It also became clear that both LP progenitors and the eye region no longer expressed *tfec* in presumptive mutant embryos (Fig. 3.3 C,D).

At 36 hpf, the observed phenotype persisted (Fig. 3.3 E-H), although there was a considerable decrease in the number of *tfec*-expressing premigratory NCCs, which was attributed to the onset of apoptosis amongst unspecified progenitors (Dutton et al. 2001). Furthermore, expression in the eyes and the developing LPs remained absent. These effects were observed in 15/51 assessed embryos over 3 independent repeats (mutants

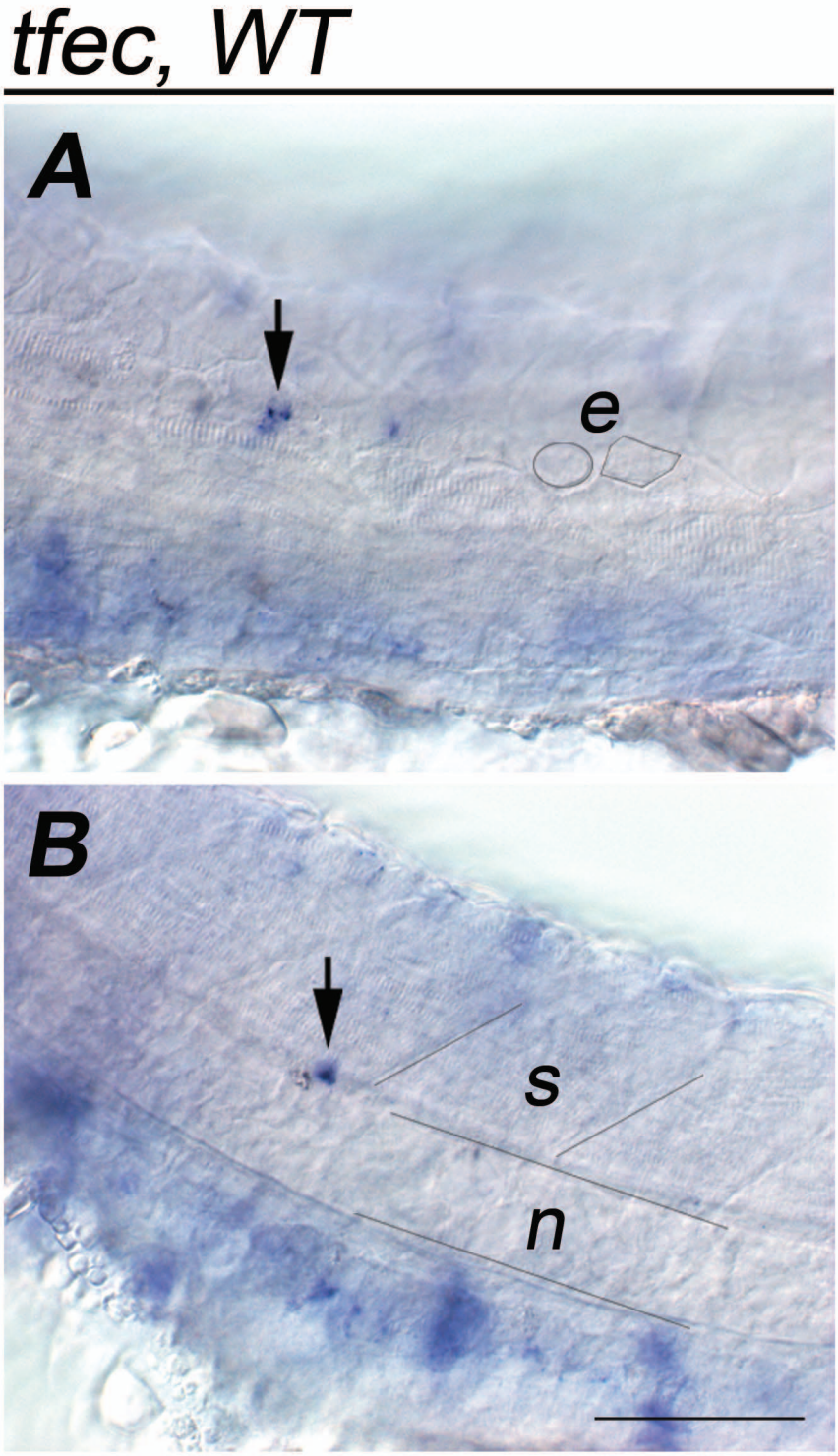


Figure 3.2. tfec is expressed in cells of the lateral and of the medial pathway at 24 hpf. *In situ hybridisation showing tfec positive cells (arrows) migrating (A) via the lateral pathway, underneath the epidermis (e; keratinocytes outlined) and (B) via the medial pathway, directly adjacent to the notochord (n; section of the tube outlined) and below the somites (s; anterior side of single muscle block outlined). Lateral views, oriented with the head to the left. tail to the right. Scale bar: 50 µm.*

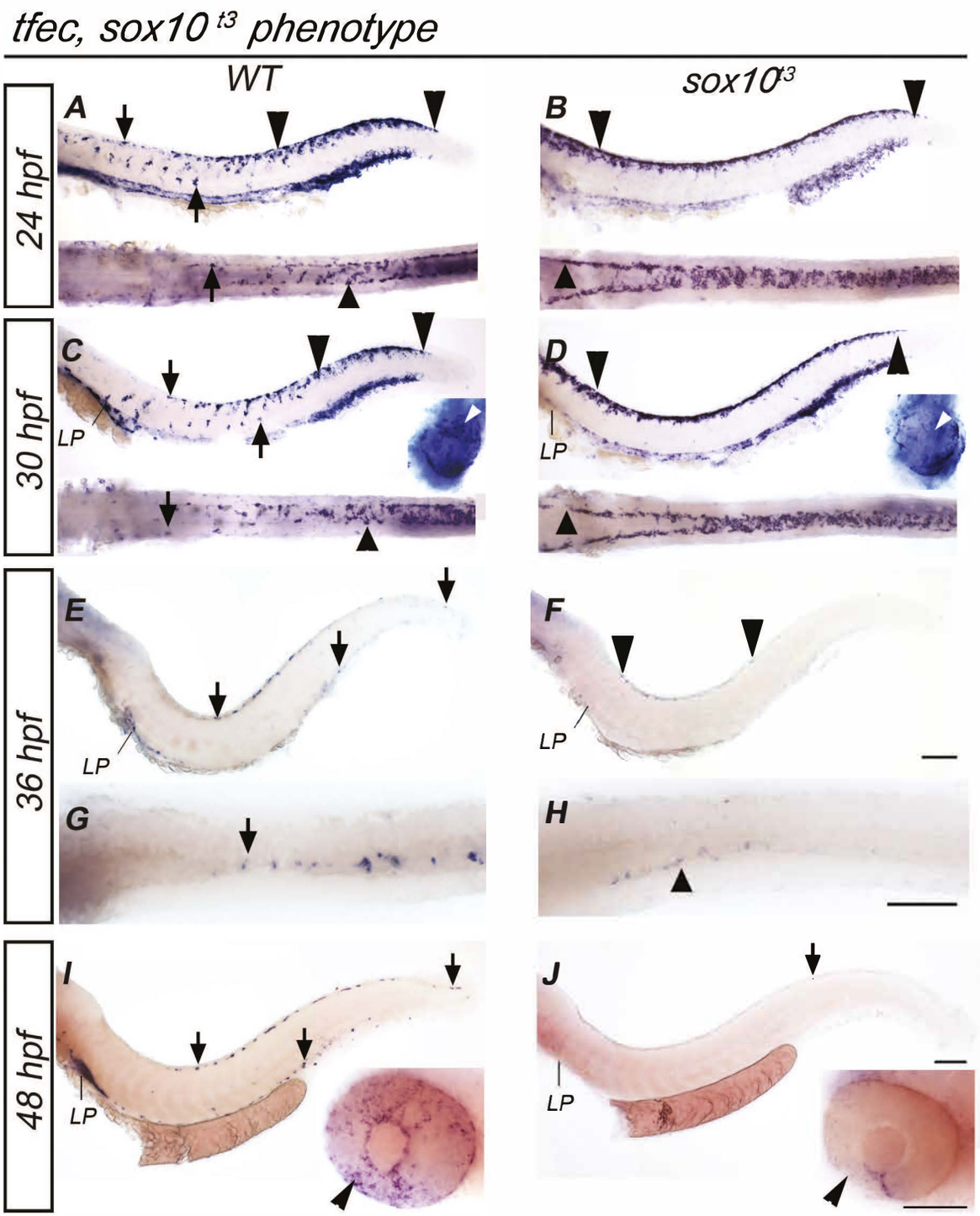


Figure 3.3. sox10 maintains tfec in iridoblasts. *In situ hybridisation results demonstrate tfec expression in WT or heterozygotes (A, C, E, G, I) and in sox10^{t3} embryos (B, D, F, H, J) at (A, B) 24 hpf. Inserts: dorsal view of the trunk. (C,D) 30 hpf. Inserts: expression in the vicinity of the RPE (white arrowheads) and dorsal view of the trunk. (E, F) 36 hpf, lateral view. (G, H) 36 hpf, dorsal view of the anterior trunk. (I, J) 48 hpf. Inserts: expression overlaying the RPE (arrowheads). From 24 hpf to 36 hpf (A-H), in presumptive sox10 mutants expression persists in the dorsally located premigratory NC domain (arrowheads), while specified iridoblasts are missing (arrows). At 36 hpf, unspecified NC progenitors in sox10 mutants (F, H) are reduced. At 48 hpf (I-J) differentiated iridophores (arrows) are absent in presumptive mutant embryos, with one escaper cell still detectable (J). Inserts show lack of tfec expression on top of the RPE (arrowheads) in sox10 mutants compared to phenotypically WT siblings. LP, lateral patches. Lateral views unless stated otherwise, oriented with the head to the left. Scale bars: 100 µm.*

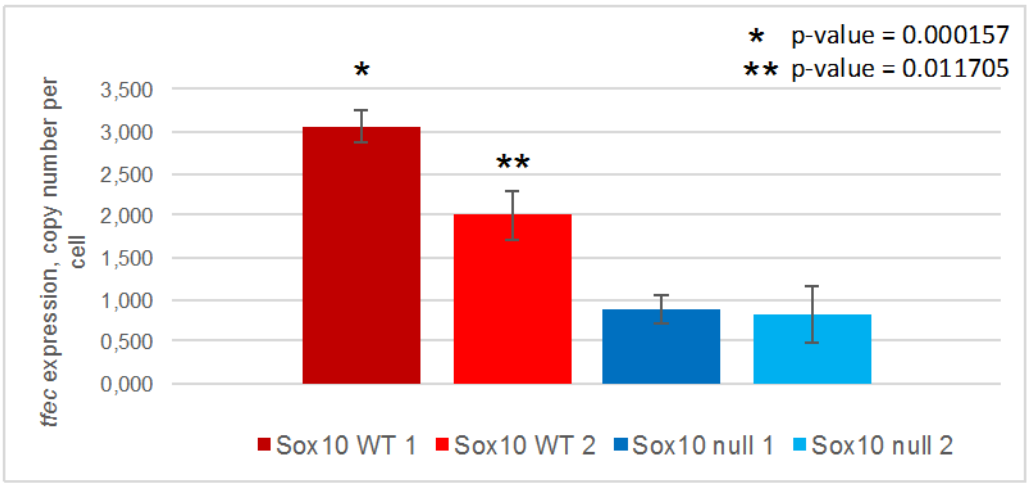


Figure 3.4. Overexpression of WT *sox10* drives ectopic upregulation of *tfec* at 6 hpf. Absolute quantification of *tfec* transcript levels via q-RT PCR. Embryos injected with WT *sox10* mRNA at the single-cell stage (two biological replicates; corrected expression per cell, red) present with a statistically significant increase in *tfec* transcript levels, compared to embryos injected with *sox10*^{m618} mRNA (corrected expression per cell, blue). Differences observed in the degree of upregulation between the biological replicates was attributed to injection of different net amounts of *sox10* WT transcript. Error bars indicate corrected standard deviation per cell, across technical replicates. Data by Dr Tatiana Subkhankulova.

according to Mendelian ratios; 0.4 < p-value < 0.5). Finally, by 48 hpf, when *tfec* expression co-localises with differentiated iridophores, 3/11 examined embryos (mutants according to Mendelian ratios; p-value > 0.95) presented with diminished *tfec* expression in all tissues. This phenotype was confirmed through 2 experimental replicates. Specifically, a very small number of escaper cells was observed in iridophore locations along the dorsal stripe and the eye, while *tfec* transcript was undetectable along the ventral and yolk sac stripes as well as on the developing LPs and swim bladder epithelium (Fig. 3.3 I,J).

These results suggested that *sox10* is important for maintenance of *tfec* expression as multipotent NCCs become specified towards the iridophore lineage. This conclusion is supported by data from overexpression studies conducted by Dr Tatiana Subkhankulova, a postdoctoral researcher in the Kelsh group. Results suggested that, following overexpression of WT *sox10* mRNA by microinjection in single-cell stage embryos, *tfec* expression levels at 6 hpf increased, compared to expression levels in siblings injected with null *sox10* mRNA (Fig. 3.4). In this experiment, absolute levels of *tfec* expression were quantified by quantitative real time PCR (q-RT PCR). Transcript copy number per cell was determined using the standard curve method.

3.2.3. *sox10* regulates *ltk* in specified iridophores

Next, *sox10* mutant embryos, fixed at the stages of interest, were used to confirm previously published data suggesting that *ltk* expression in iridoblasts depends upon *sox10* function (Lopes et al. 2008).

Detecting *ltk* expression by *in situ* hybridisation proved challenging during initial experiments, due to increased background: signal ratio, as well as contamination of plasmid stocks used to generate *ltk* RNA probes with unknown vectors. These problems were solved by:

1) Re-synthesising the probe using primers against *ltk* cDNA previously designed by S. Lopes (Appendix Table II.3). The cloning procedure was repeated to generate a new plasmid vector from which the RNA probe could be generated. This step ensured that the plasmid carried the desired sequence.
2) Hydrolysing the DIG-labelled RNA probe (refer to materials and methods section 2.5.1) so that its size was reduced from 1.2 kb to 600 nucleotides. This resulted in increased tissue permeability, and therefore effective binding to mRNA molecules, especially within cells located in non-superficial tissues.
3) Pre-adsorbing the probe (refer to materials and methods section 2.5.1) to increase specificity.

4) Enhancing stringency by increasing the hybridisation temperature (refer to materials and methods section 2.5.3). Whilst the rest of the probes were hybridised at 67.5 ºC, the *ltk* probe was hybridised at temperatures between 69 ºC to 70 ºC. During the second day, stringency washes were carried out at 67.5 ºC. This treatment significantly reduced non-specific hybridisation of the probe to targets other than *ltk*.

Expression of *ltk* was assessed in *sox10* mutants and in their WT siblings at 18 hpf, 24 hpf, 30 hpf, 36 hpf and 60 hpf. All results presented here were derived from one independent repeat for each developmental stage, and were compared to previously published results (Lopes et al. 2008). At 18 hpf, *ltk* expression was unaffected in 100% of assessed embryos (N ≈ 25). At this stage, the transcript was not yet detectable within NCCs of the trunk, although expression was observed in the cNC domain and in cells inside the notochord (Fig. 3.5 A,B). At 24 hpf (N ≈ 25), presumptive WT embryos displayed the previously established *ltk* expression pattern (Lopes et al. 2008). In these embryos, a subpopulation of premigratory NCCs at the posterior trunk of the embryo, as well as medially migrating cells towards the ventral trunk, expressed *ltk* (Fig. 3.5 C,D,E). Expression in the developing LPs was also detectable at this stage (data not shown). In these experiments, the increased level of background signal in the head did not allow assessment of *ltk* expression in the RPE region (data not shown). In accordance with previously published data (Lopes et al. 2008), presumptive mutant embryos showed significant expansion of the *ltk*-positive premigratory NC domain and a lack of migrating cells (Fig. 3.5 F,G,H). In these embryos, expression in LP progenitors appeared reduced or absent (data not shown).

At 30 hpf (N ≈ 25), in presumptive WT embryos *ltk* expression marked positions occupied by specified iridoblasts along the dorsal trunk and the medial migratory pathway (Fig. 3.5 I,J,K). *ltk* positive cells were bilaterally distributed on the dorsal side of the tail, indicating the presence of premigratory NCCs at this location (Fig. 3.5 I). Expression was seen associated with the RPE and in the developing LPs (data not shown). In likely *sox10* mutants, *ltk* expression was eliminated along migratory pathways while the tNC domain was expanded (Fig. 3.5 L,M,N). In these embryos, staining was reduced in the vicinity of the RPE and in LP progenitors (data not shown). Similar WT and mutant phenotypes were observed in 36 hpf embryos (N ≈ 25) (Fig. 3.6 A-D).

By 60 hpf (N ≈ 25), in presumptive WT embryos *ltk* expression labelled differentiated iridophores along the dorsal, ventral and yolk sac stripes (Fig. 3.6 E,F). The gene was also expressed in the layer overlaying the RPE (data not shown), the developing LPs and the swim bladder epithelium (Fig. 3.6 E). In *sox10* mutants, all expression of *ltk* was largely eliminated, indicating lack of differentiated iridophores, in line with what is known

ltk, sox10 t3 phenotype

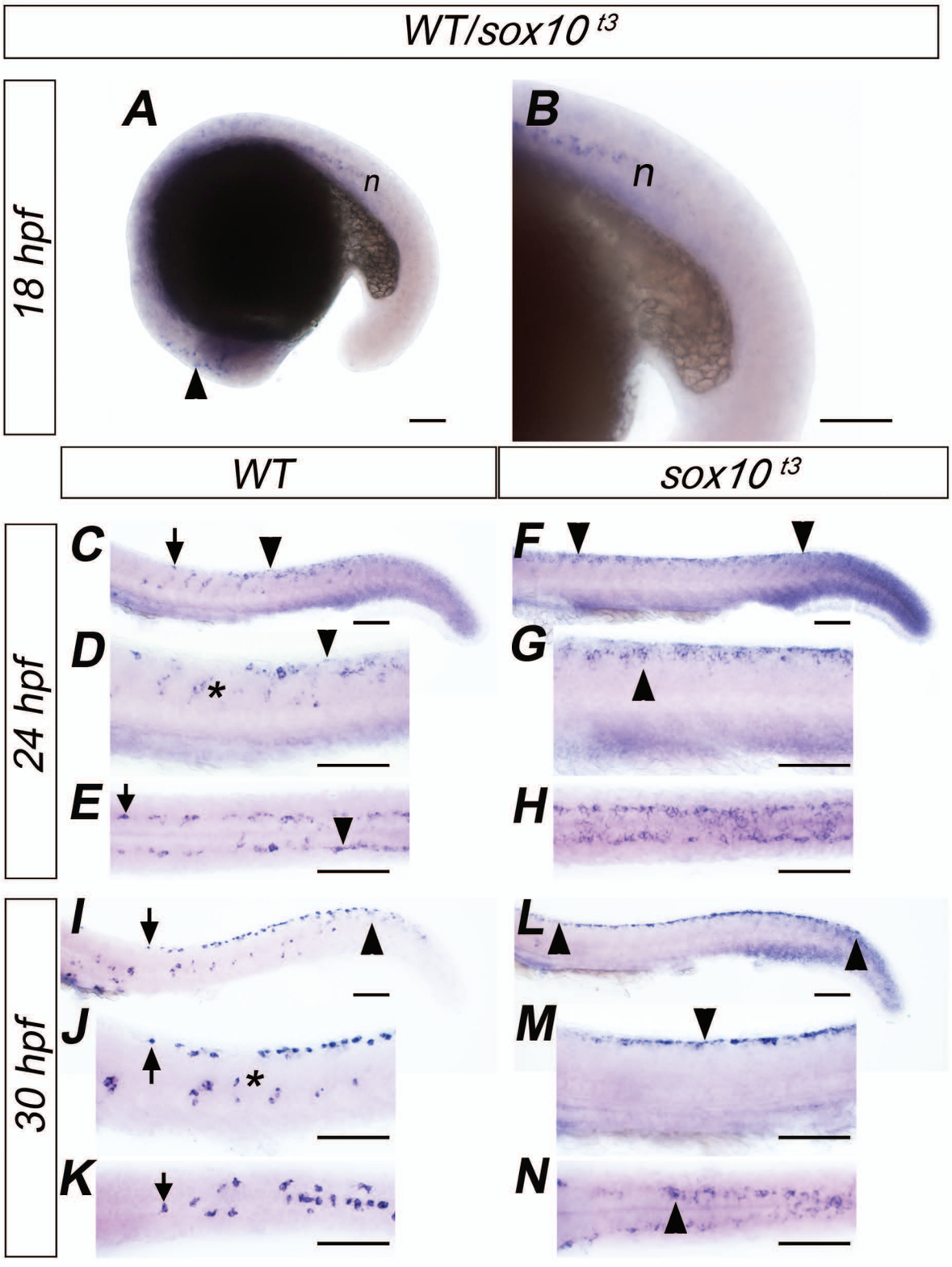


Fig. 3.5

Figure 3.5. *sox10* upregulates *ltk* in iridoblasts. *In situ* hybridisation results demonstrate *ltk* expression in WT or heterozygotes (A, B, C-E, I-K) and in *sox10*^{t3} embryos (A, B, F-H, L-N). At 18 hpf (A, B) all embryos show the WT phenotype, with the transcript present in cNCCs (arrowhead) and in the notochord. At 24 hpf and at 30 hpf, in WT or heterozygous embryos (C-E, I-K) *ltk* is expressed premigratory NCCs (arrowheads) and specified iridoblasts along the dorsal trunk (arrows) and the migratory pathways (asterisks). Presumptive mutants (F-H, L-N) present with expansion of the premigratory NC domain (arrowheads) and absence of specified iridoblasts. A-D, F, G, I, J, L, M: lateral views. E, H, K, N: dorsal views. Oriented with the head to the left. Scale bars: 100 μm.

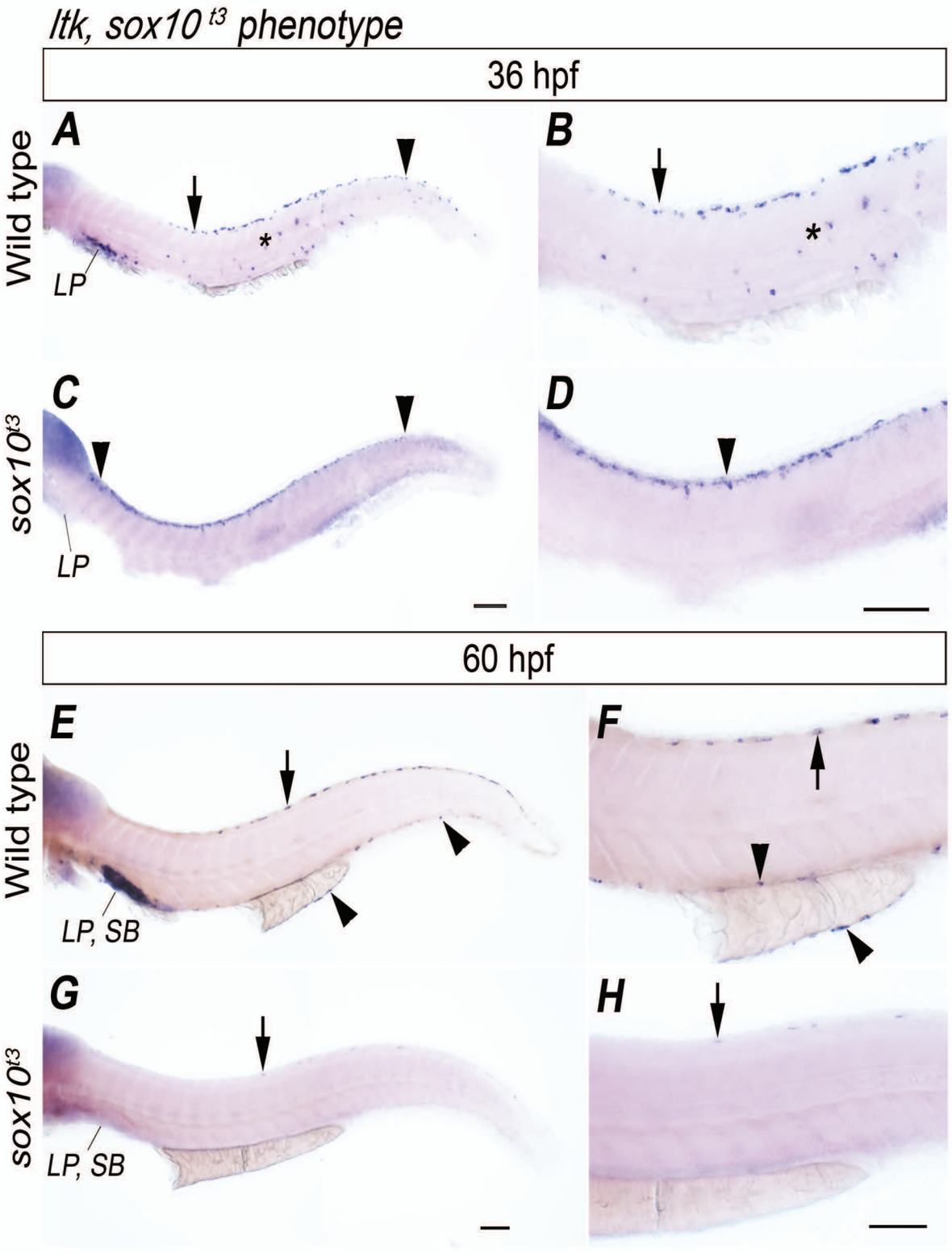


Figure 3.6. In sox10 mutants, unspecified NCCs expressing ltk fail to generate iridophores. *In situ hybridisation results demonstrate ltk expression in WT or heterozygotes (A, B, E, F) and in sox10^{t3} embryos (C, D, G, H). At 36 hpf, ltk expression is normally seen in specified iridoblasts which are dorsally positioned (arrows), ventrally migrating (asterisks) and located in the developing LPs, as well as in premigratory NCCs in the tail region (arrowhead) (A, B). In sox10 mutants, specified iridoblasts are absent and the premigratory NC domain is expanded towards the anterior (C, D). At 60 hpf ltk is normally expressed in mature iridophore locations along the dorsal (arrows), ventral and yolk sac stripe (arrowheads), in the LPs and the swim bladder epithelium (E, F). In sox10 mutants, expression is only detectable in escaper iridophores along the dorsal trunk (arrows) (G, H). LP, lateral patches; SB, swim bladder. Lateral views, oriented with the head to the left. Scale bars: 100 μm.*

about the *sox10* mutant phenotype (Lopes et al. 2008). Only a small number of cells expressing *ltk* was observed along the dorsal trunk (Fig. 3.6 H), corresponding to escaper iridophores (Lopes et al. 2008).

3.2.4. Knocking-down *tfec* results in lack of iridophores

To assess the effects of *tfec* loss of function on potential downstream targets, the gene was knocked-down using morpholino technology. The morpholino sequence was designed by Dr J. Lister (Virginia Commonwealth University, USA) as a splice site blocker. The boundary between exon 5 and intron 5 of the *tfec* sequence was targeted, hence the morpholino is referred to as 'E5I5'. Morpholinos were injected into the yolk sac of the embryo at the single-cell stage. A series of injection dosages were tested to achieve efficient knockdown with the least possible phenotypic variability within injected embryos combined with low death rates and settled for injecting 4.6-6.0 ng of morpholino per embryo.

Following optimisation of the injection technique, WT (WIK) embryos were injected with 4.6-6.0 ng of E5I5 morpholino. At 3 dpf, almost 100% of injected embryos presented with no or with very few iridophores, as was previously shown (J. Lister, unpublished data) (Fig. 3.7). The experiment was repeated on *mitfa*^{w2} mutant embryos (*nacre*) to confirm that the morpholino injection had the expected outcome, according to the previously described phenotypes (J. Lister, unpublished data). The predicted effect was that *nacre* embryos, which inherently lack melanophores, lose both iridophores and xanthophores upon morpholino injection. The phenotype was, indeed, replicated. The striking loss of iridophores from both the dorsal and the ventral stripes was quantified (Fig. 3.8), but the loss of xanthophores is not shown in this study.

Phenol red was added to the morpholino-containing solution to be injected. To confirm that the observed phenotypes were not due to injection artefacts, the same concentration of phenol-red without morpholino was injected in siblings. Phenol-red injected siblings were screened and lack of any iridophore phenotypes was confirmed (data not shown).

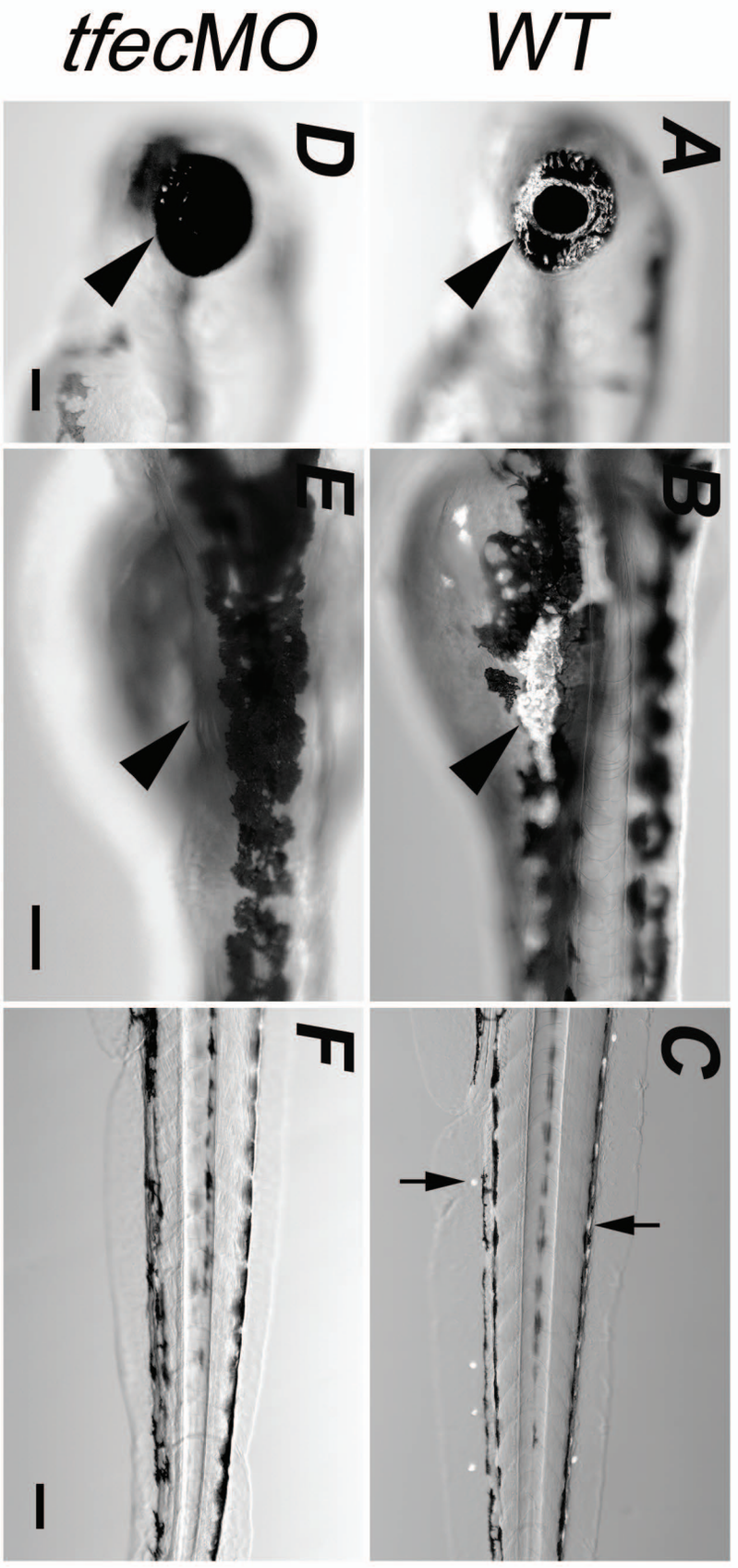


Figure 3.7. Knockdown of tfec in WT (WIK) embryos results in loss of iridophores. *Live imaging of uninjected (WT) embryo (A-C) and representative sibling injected with tfec MO (D-F) at 2 dpf under incident light. Iridophores are absent from the eyes (A, D; arrowheads), lateral patches (B, F; arrowheads) and both the dorsal and ventral stripe of the trunk (C, F; arrows) of injected embryos. Lateral views, oriented with the head to the left. Scale bars: 100 µm.*

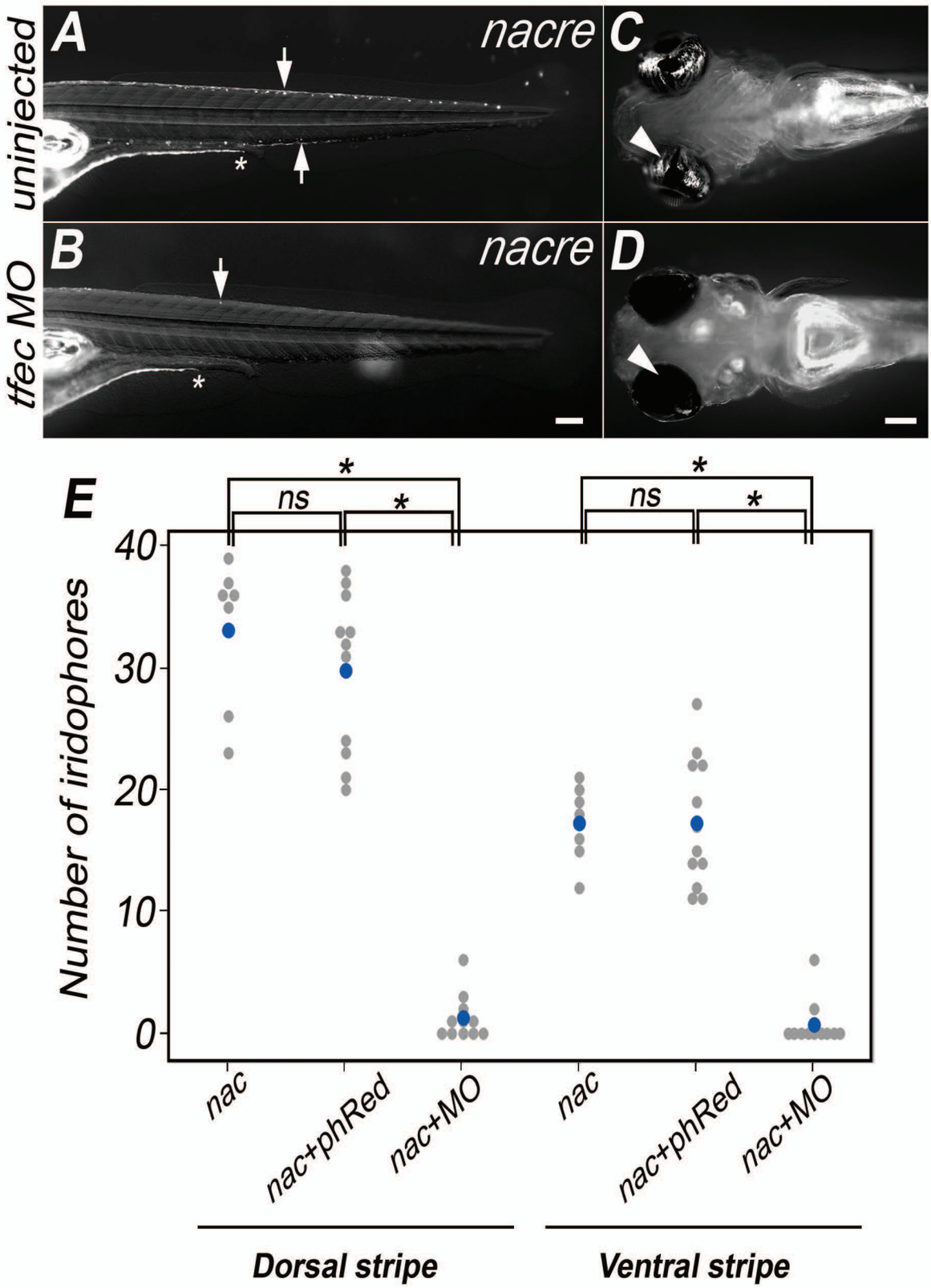


Figure 3.8. Knockdown of tfec in mitfa^{w2} (nacre) embryos results in loss of iridophores at 5 dpf. *Live imaging of uninjected (nacre) embryo (A, C) and representative sibling injected with tfec MO (B, D) under incident light. In injected embryos, iridophores are absent from the eyes (C, D; arrowheads), the dorsal and ventral stripe of the trunk (A, B; arrows), while there is a reduction is cells along the yolk sac stripe, which appears less extended (A, B; asterisks). (E) Quantitation of iridophore number on the dorsal and ventral stripes of uninjected, phenol red-injected and morpholino-injected embryos at 5 dpf. Blue dots represent mean values. * p-values < 0.001; ns, not significant. A, B: lateral views. C, D: dorsal views. Oriented with the head to the left. Scale bars: 100 μm*

3.2.5. *tfec* and *ltk* form a positive feedback loop

The next step was to identify relationships between *tfec* and *ltk*. First, *tfec* expression was assessed in *ltk*^{ty82} mutants and WT siblings at the stages of interest using *in situ* hybridisation. As described in section 3.2.2, the Pearson's χ^2 test was used to interpret results obtained from *in situ* hybridisation experiments (refer to section 2.8.1). The null hupothesis stated that, in a sample of mixed WT, heterozygous and homozygous mutant embryos, a defined alternative phenotype matched the expected proportion of homozygous mutants (1 mutant: 3 WT phenotype). As before, for p-values larger than 0.1 the null hypothesis was accepted, and the conclusion that the observed alternative phenotype was an effect of the mutation was supported.

At 18 hpf, 100% of assessed embryos showed the WT *tfec* expression pattern (N = 79; 4 independent repeats; p-value < 0.001). Similarly, no statistically significant deviations from the WT were observed amongst pools of WT and *ltk* mutant embryos at 24 hpf. Over 5 independent repeats, 194 embryos at 24 hpf were examined in total. Of these, only 5 individuals (p-value < 0.001) presented with observable expansion of the NC domain, an effect attributed to staging variation within the samples. All remaining embryos displayed the WT phenotype (data not shown). Thus, in *ltk* homozygous mutants effects do not manifest concomitantly to those in *sox10* mutants, indicating different roles for the two genes during iridophore development.

At 30 hpf, 3 independent experiments revealed a consistent and statistically significant decrease in the number of *tfec* positive cells along the dorsal trunk of potential *ltk* homozygous mutants compared to their WT siblings (effect observed in 23/80 assessed embryos; mutants according to Mendelian ratios; 0.3 < p-value < 0.5) (Fig. 3.9 A-D). The numbers of *tfec*-expressing cells on the dorsal trunk were scored in presumptive WT and mutant embryos and the two-sample t-test was used to establish statistical significance of the deviations between the sample means (Fig. 3.9 I). Additionally, the number of cells in both the ventrally migrating cells and the eyes appeared reduced in the affected embryos, while LP progenitors and the posteriorly detectable premigratory NC domain were not observably affected (Fig. 3.9 B,D).

At 36 hpf, the described phenotype persisted in 28/102 assessed embryos, examined over 4 independent experimental repeats (mutants according to Mendelian ratios; 0.5 < p-value < 0.7) (Fig. 3.9 E-H,I). Quantification of *tfec*-positive cells along the dorsal trunk did not suggest an added decrease compared to affected 30 hpf embryos. Instead, the observably increased number of cells present in figure 3.9 (F, H), indicated either that experimental conditions increased sensitivity of transcript detection in this sample, or that a different, unidentified type of precursor expressed *tfec* at this stage in addition to

the iridoblast. Alternatively, the result may be interpreted as a delay in *tfec* activation in the iridoblast, although this is unlikely considering the lack of mature iridophores present in *ltk* mutants. The effect of loss of *ltk* function on the *tfec*-expressing region of the eye was again observed, although the LP regions retained *tfec*-positive cells.

By 48 hpf, 44/60 assessed embryos (WT according to Mendelian ratios; 3 independent experimental repeats; $0.7 < \text{p-value} < 0.8$) displayed between 20 and 30 *tfec* positive cells in iridophore positions along the dorsal trunk. The remaining embryos had between 0-4 escapers along the dorsal trunk, 0-5 escapers along the ventral trunk, few cells remaining in the LP domain and no swim bladder staining was detectable (data not shown). Together, these data suggested that, although loss of *ltk* function did not elicit an effect in the initial specification of iridoblasts, the gene is crucial for survival of specified iridoblasts, since maintenance of *tfec* expression in the specified lineage requires *ltk* function.

To investigate whether *tfec* upregulates *ltk* in the iridophore lineage, embryos injected with *tfec* morpholino, dissolved in water and mixed with phenol red dye to visualise the solution, were fixed and *in situ* hybridisation was performed. *ltk* expression was assessed at 30 hpf and at 3 dpf in *tfec* morphants, versus phenol red-injected siblings as negative controls. At 30 hpf, there was a significant reduction in the number of *ltk*-positive cells in the dorsal trunk of a subset of embryos injected with *tfec* MO, while in phenol red-injected siblings *ltk* expression strongly labelled specified iridoblasts of the dorsal trunk (Fig. 3.10 A-D). At 3 days post-fertilisation (dpf), morphants lacked the majority of iridophores of the dorsal and of the ventral trunk, while *ltk* expression was maintained in the swim bladder epithelium (Fig. 3.10 E-H).

Together, these data supported the existence of a positive feedback loop between *tfec* and *ltk*. Data presented in sections 3.2.2 and 3.2.3 suggested that both *tfec* and *ltk* expression in specified iridoblasts depended upon *sox10* function. Therefore, it was hypothesised that *sox10*-dependent maintenance of *tfec* expression initiated specification of iridoblasts and triggered the establishment of the *tfec*/*ltk* feedback loop crucial for their survival.

tfec, ltk^{ty82} *phenotype*

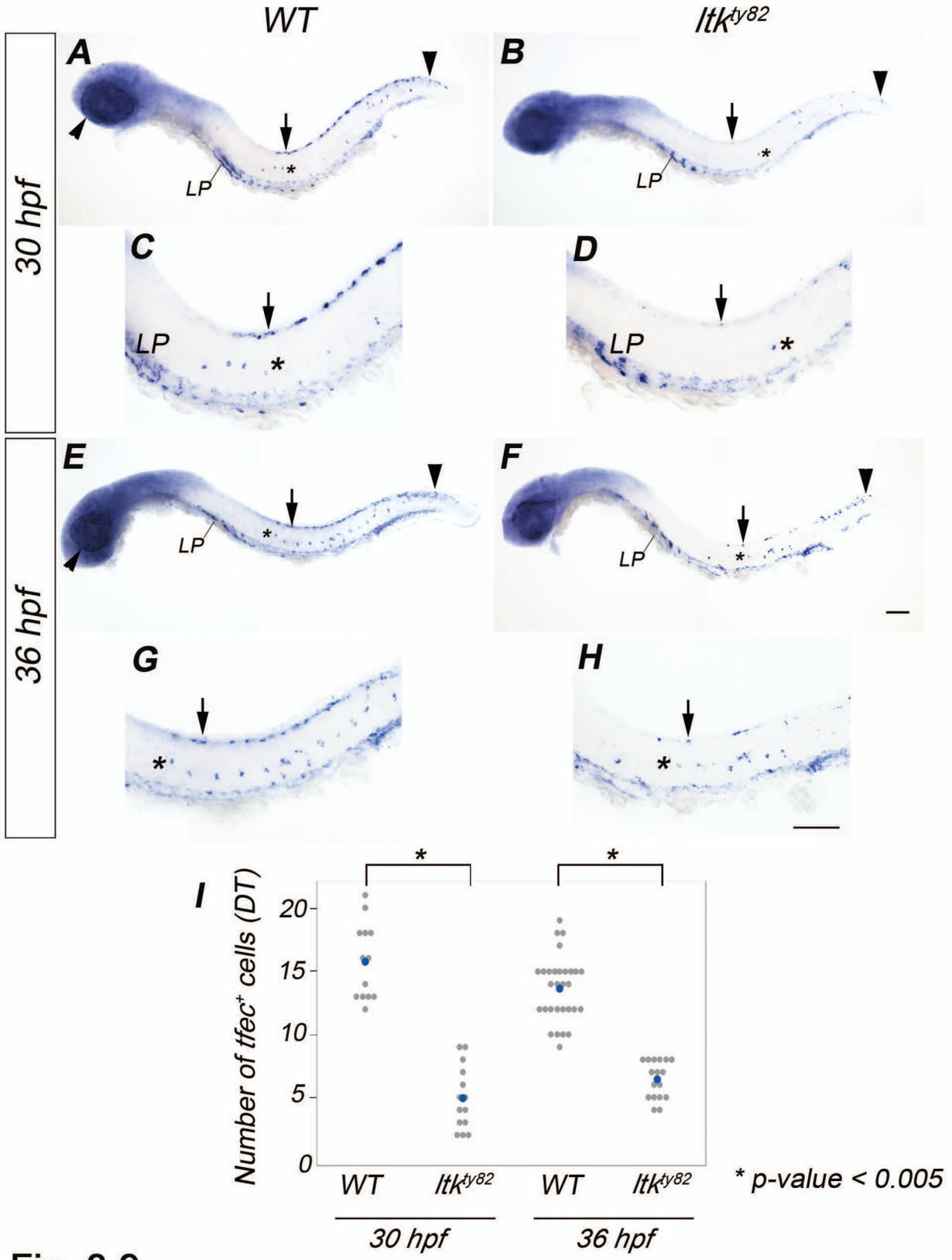


Fig. 3.9

Figure 3.9. *ltk* upregulates *tfec* in iridoblasts. *In situ* hybridisation results demonstrate *tfec* expression in WT or heterozygotes (A, C, E, G) and in presumptive *ltk*^{ty82} embryos (B, D, F, H). At both 30 hpf and 36 hpf, WT *tfec* expression is observed in dorsally located (arrows) and migrating (asterisks) iridoblasts, premigratory NCCs (vertical arrowheads), developing LPs and overlaying the RPE (angled arrowheads) (A, C, E, G). At both stages, presumptive *ltk* mutants (B, D, F, H) show reduction of *tfec* positive cells along the dorsal trunk (arrows) and the migrating pathways (asterisk), while premigratory NCCs in the dorsal tail are still detectable (arrowheads). The LPs are not observably affected in these embryos. (I) Quantitation of *tfec* positive iridoblasts at the dorsal trunk of presumed WT and mutant embryos at both stages. * p-value < 0.001. LP, lateral patches; DT, dorsal trunk. Lateral views, oriented with the head to the left. Scale bars: 100 μm.

3.2.6. Both *tfec* and *mitfa* regulate *pnp4a* expression

It has been previously reported that *pnp4a* is expressed in specified iridoblasts, thus serving as an early iridophore marker (Curran et al. 2010). By performing *in situ* hybridisation on embryos fixed at each of the stages of interest, it was demonstrated that *pnp4a* expression during early stages of NCC specification was not only detectable in iridoblasts, but also in other NC-derived migrating cell types, as well as in premigratory NCCs, during early stages of specification (Fig. 3.11).

pnp4a expression was first detectable at 18 hpf in very few cells of the cNC domain and in association with the presumptive RPE (Fig. 3.15 I-J). At 24 hpf, *pnp4a* was expressed strongly in the vicinity of the RPE, as well as in regions of the forebrain that likely correspond to NC derivatives, based on data showing *mitfa* expression in these regions (Fig. 3.11 A,B,F,G; Fig. 4.10). Expression was prominent in premigratory cNCCs and tNCCs as well as in ventrally migrating cells of both the lateral and the medial pathway (Fig. 3.11 B,F,G). By 30 hpf, expression in the brain had significantly subsided, but was strongly maintained in the eyes and in mid to posterior dorsal trunk regions, presumed to correspond to specified iridoblasts (Fig. 3.11 C,D). *pnp4a* was found to be weakly expressed in a number of cells of the anterior dorsal trunk domain and of the ventral hindbrain, where iridoblasts do not localise (Fig. 3.11 C,D) and remained detectable in the dorsally located premigratory NC region of the tail, as well as in laterally and medially migrating cells (Fig. 3.11 C,H,I). Although the expression pattern in the above stages strikingly resembled the melanoblast-specific *mitfa* expression pattern at the corresponding stages (J A Lister et al. 1999), by 48 hpf, *pnp4a* expression resembled the iridophore pattern, as previously reported (Curran et al. 2010) (Fig. 3.11 E).

These results indicated that *pnp4a* was not solely an iridophore marker, but was also strongly upregulated in other NC derivatives at stages prior to iridophore differentiation. Based on the striking resemblance of the *pnp4a* expression pattern to that of *mitfa*, a potential relationship between these two genes was investigated.

pnp4a expression was assessed in *mitfa*^{w2} mutants by *in situ* hybridisation on each of the stages of interest. No observable alterations in *pnp4a* expression were observed in batches of WT/*mitfa*$^{+/-}$/*mitfa*$^{-/-}$ embryos at 18 hpf (N ≈ 25; one experimental replicate) (data not shown). At 24 hpf, a severe reduction in *pnp4a* expressing cells was observed in 11/49 examined embryos (mutants according to Mendelian ratios; 2 independent experimental repeats; $0.5 < \text{p-value} < 0.7$) (Fig. 3.12 A-D). The number of *pnp4a* positive cells of the premigratory NC domain was reduced to an average of 4 cells, whereas an average of 3 cells were seen on the migratory pathways (n = 7 embryos scored). Expression was eliminated in emerging LP progenitors and appeared reduced in the head, where only a few cells still expressed *pnp4a* (data not shown).

At 30 hpf, a reduction in *pnp4a* positive cells was again observed in 11/31 embryos (mutants according to Mendelian ratios; 2 independent experiments; 0.1 < p-value < 0.2). The effect was prominent in LP progenitor regions, along the migratory pathways and the ventral trunk, where cells positioned in a *mitfa*-like pattern were missing (Fig. 3.12 E-H). The distribution of cells maintained along the dorsal trunk and migratory pathways at this point was strongly reminiscent of *tfec* expression in specified iridoblasts at these locations in corresponding stages (Fig. 3.1 C).

At 36 hpf, a different phenotype was demonstrated as a likely result of loss of *mitfa* function. The WT *pnp4a* expression pattern normally featured a noticeable reduction in the number of cells that resembled the *mitfa* expression pattern. Instead, the pattern at this stage was similar to that of *tfec* expression, labelling interspersed cells along the dorsal trunk and a small number of migrating cells towards the posterior trunk and tail region (Fig. 3.12 I,K). Over 2 independent experimental repeats, in examined samples of mixed WT/*mitfa*$^{+/-}$/*mitfa*$^{-/-}$ embryos, 12/46 individuals (mutants according to Mendelian ratios; p-value = 0.9) displayed a striking increase in the number of migrating cells compared to their siblings, which showed the expected *pnp4a* expression pattern (Fig. 3.12 I-L). This increase parallels the known effect of an increased number of iridophores in *mitfa*^{w2} mutants (J A Lister et al. 1999). The LP regions appeared identical amongst embryos in these samples.

Overall, the results described above suggested that *pnp4a* expression depends primarily on *mitfa* between 24 and 30 hpf, however, by 36 hpf upregulation only occurs in an iridophore-specific manner. The role of *tfec* in *pnp4a* regulation was then investigated, since the former is presumed to be the master regulator of iridophore development.

In situ hybridisation experiments were performed on fixed WT embryos and *tfec* morphant siblings. While all of the examined uninjected WT embryos displayed the previously described *pnp4a* expression pattern at 24 hpf (data not shown; Fig. 3.11 A,B,F,G), the observed phenotypes amongst morpholino-injected siblings (N = 22) featured different degrees of reduction in the numbers of *pnp4a* positive cells either residing along the premigratory NC domain, or migrating towards the ventral trunk (Fig. 3.13). The most severe phenotype observed involved a complete lack of *pnp4a* expressing cells in both domains (Fig. 3.13 D). It was concluded that *tfec* is essential for driving *pnp4a* expression from early stages of development, however it was not possible to draw assumptions on whether loss of *tfec* function leads to abolishing of *mitfa*-dependent expression based on knock-down data. This was because morpholino-mediated loss of function is associated with a high degree of variability and often with non-specific effects related to induction of apoptosis or to unintended targeting of other genes (Schulte-Merker and Stainier 2014).

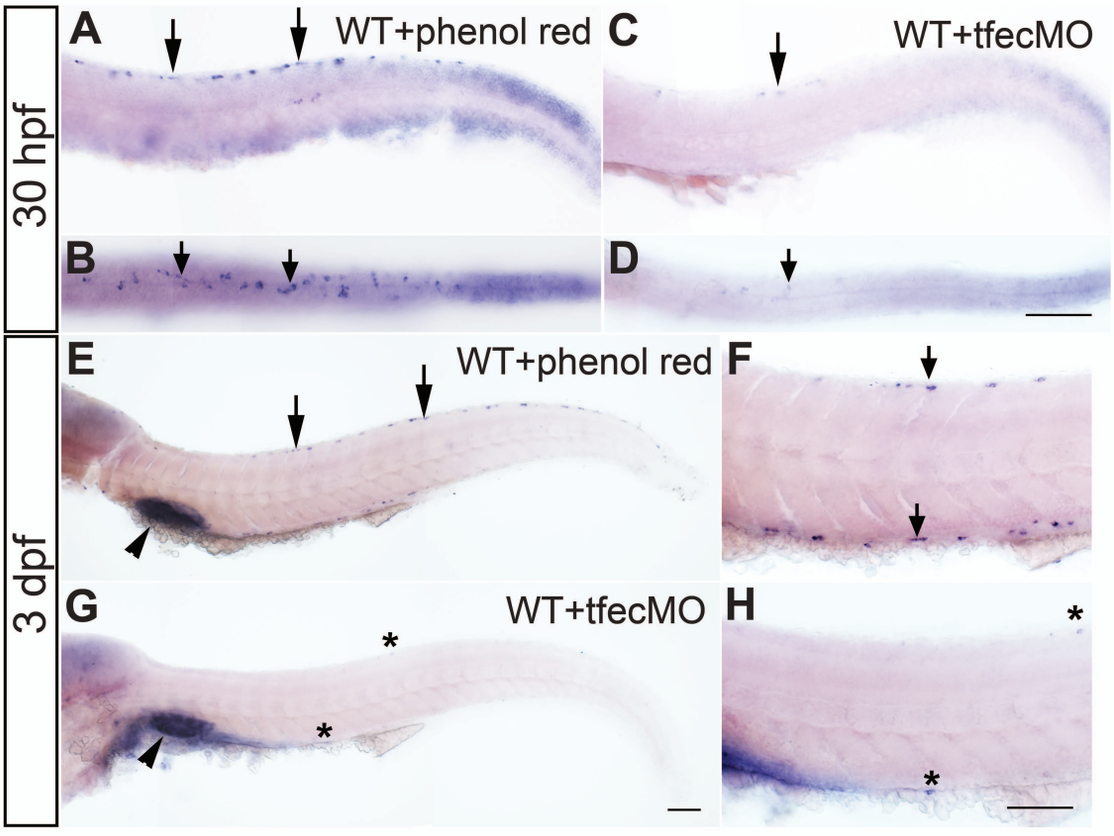


Figure 3.10. *tfec* upregulates *ltk* in iridoblasts. *In situ* hybridisation results demonstrate *ltk* expression in phenol red injected embryos (A, B, E, F) and in *tfec* morphant siblings (C, D, G, H). At 30 hpf, in phenol red injected embryos (A,B) *ltk* is expressed in specified iridoblasts (arrows), which are largely absent in *tfec* morphants (C,D). At 3 dpf, mature iridophores along the dorsal and ventral stripes (arrows) and on the LPs (arrowheads), as well as cells lining the swim bladder (arrowheads) express *tfec* in phenol red injected embryos (E, F). *tfec* morphants maintain *ltk* expression in few escaper iridophores (asterisks) and in cells lining the swim bladder (arrowhead) (G, H). A, C, E-H: lateral views. B, D: dorsal views. Oriented with the head to the left. Scale bars: 100 µm.

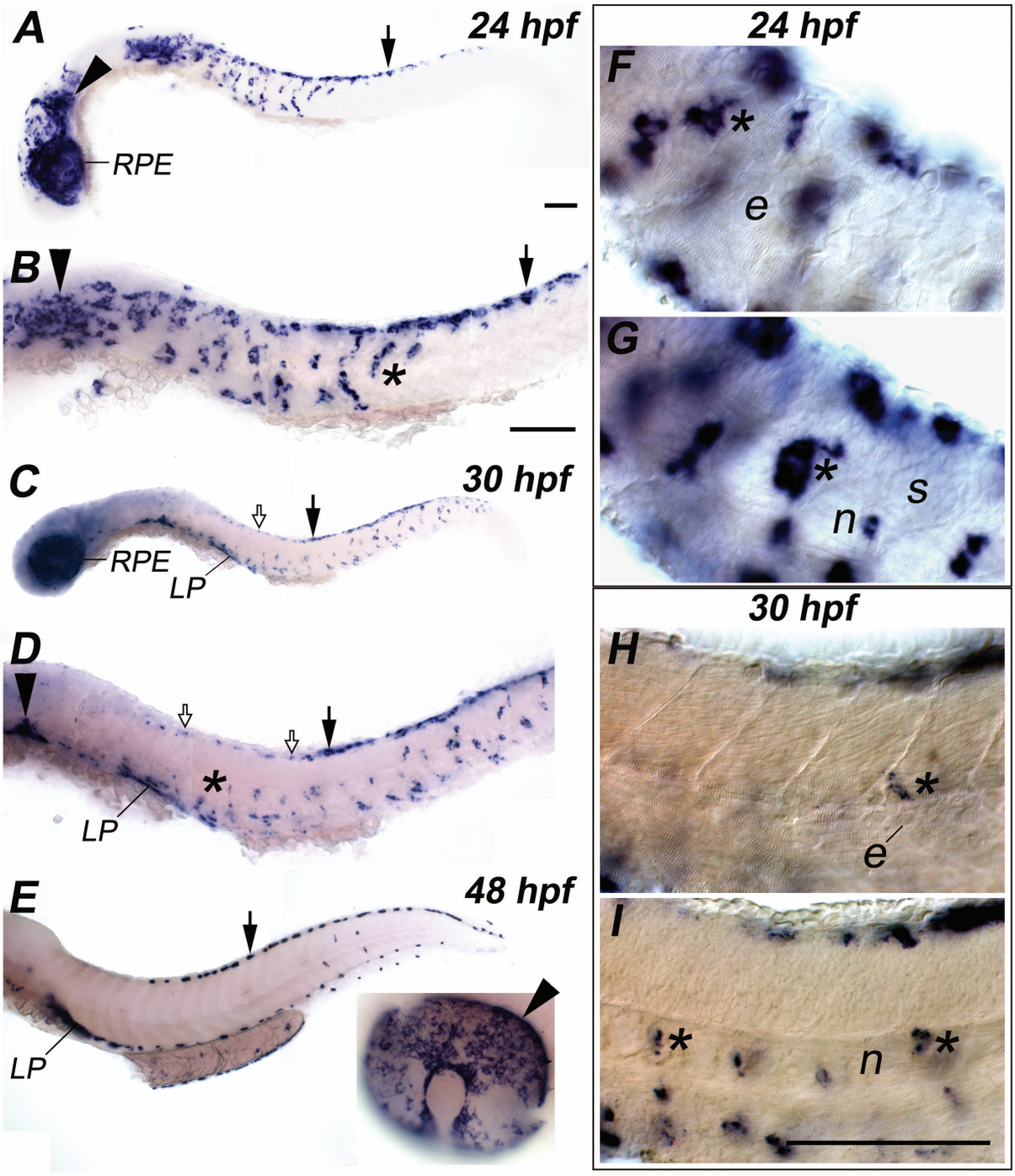


Figure 3.11. pnp4a is expressed during early specification of diverse NC derivatives and is a marker of differentiated iridophores. *In situ hybridisation results show that in WT embryos at 24 hpf (A, B, F, G), pnp4a is expressed in the premigratory tNC domain (arrows), in cNCCs (arrowheads), in migrating NCCs (asterisk) along the lateral (F) and the medial (G) pathways and in the vicinity of the RPE. Cells on the lateral pathway are seen migrating below the epidermis (F) and on the medial pathway between the somite blocks and the notochord (G). At 30 hpf (C, D), the RPE region, developing LPs, dorsally located iridoblasts (solid arrows) and migrating NC derivatives (asterisk) strongly express pnp4a. More anteriorly located NC derivatives appear to weakly express pnp4a (empty arrows). The arrowhead (D) indicates a stained region of the ventral hindbrain. (H) and (I) show expression in laterally and medially migrating cells (asterisks), respectively, at 30 hpf. (E) demonstrates expression in differentiated iridophore positions of the trunk (arrow), the layer overlaying the RPE (insert, arrowhead) and LPs at 48 hpf. LP, lateral patches; RPE, retinal pigment epithelium; e, epidermis; n, notochord; s, somites. Lateral views, oriented with the head to the left. Scale bars (A-E): 100 µm, (F-I): 50 µm.*

pnp4a, mitfa^{w2} phenotype

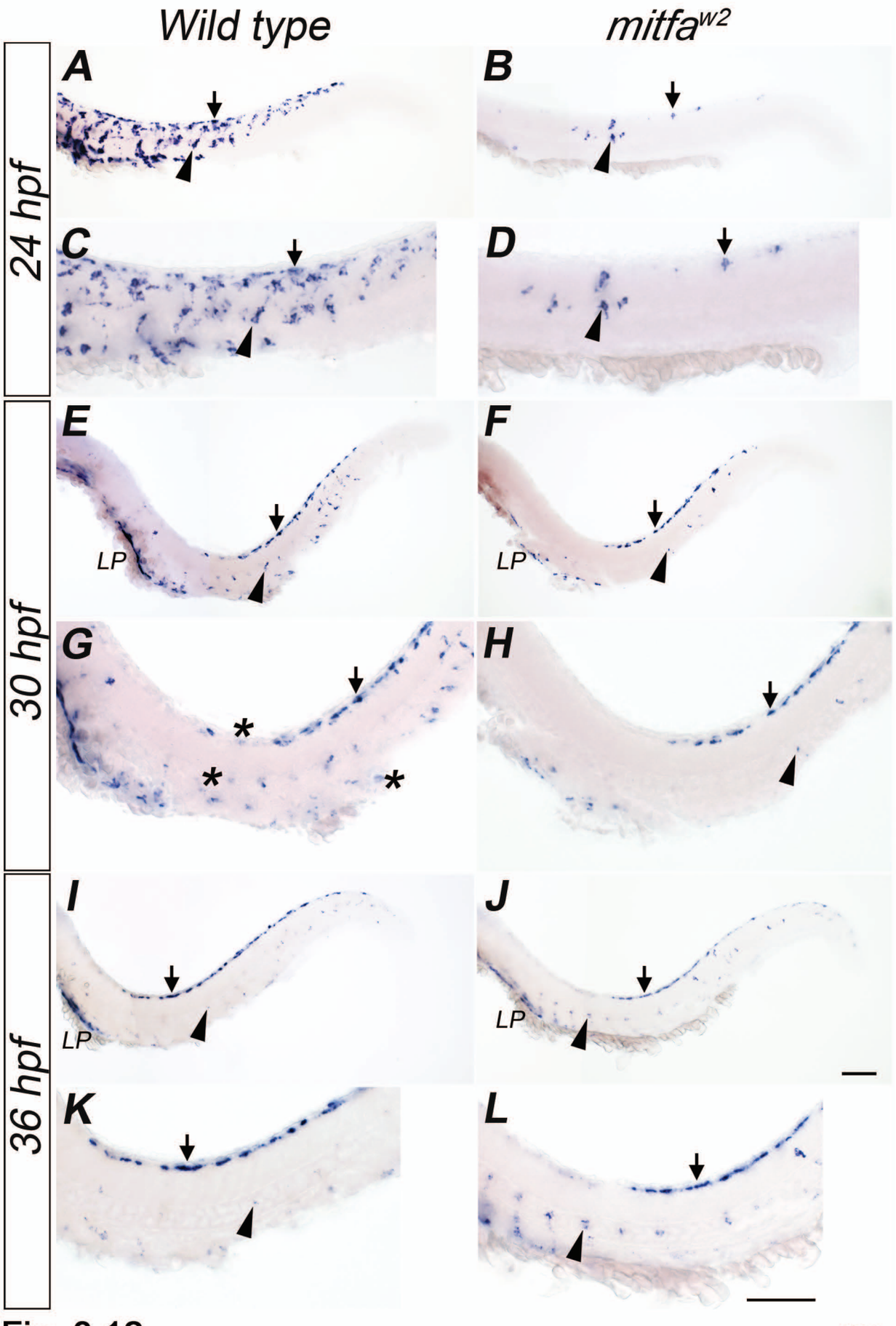


Fig. 3.12

Figure 3.12. *mitfa* upregulates *pnp4a* at 24 hpf, but this effect is gradually diminished. *In situ* hybridisation results show *pnp4a* expression in WT or heterozygotes (A, C, E, G, I, K) and in presumptive *mitfa*^{w2} embryos (B, D, F, H, J, L). At 24 hpf (A-D, lateral views), a striking number of *pnp4a* positive cells are missing in mutant embryos, both from premigratory NC regions (arrows) and from the migratory pathways (arrowheads). At 30 hpf (E-H, lateral views), loss of *mitfa* function does not affect specified iridoblasts along the dorsal trunk (arrows), in the migratory pathways (arrowheads) and in the developing LPs, whereas *mitfa*-dependent *pnp4a* expression in the dorsal and ventral trunk and in migratory pathways is lost (asterisks). At 36 hpf (I-L), an increase in migrating cells was observed in presumed mutants (arrowheads), while dorsally located iridoblasts (arrows) and the developing LPs maintain expression in all embryos. Lateral views, oriented with the head to the left. Scale bars: 100 μm.

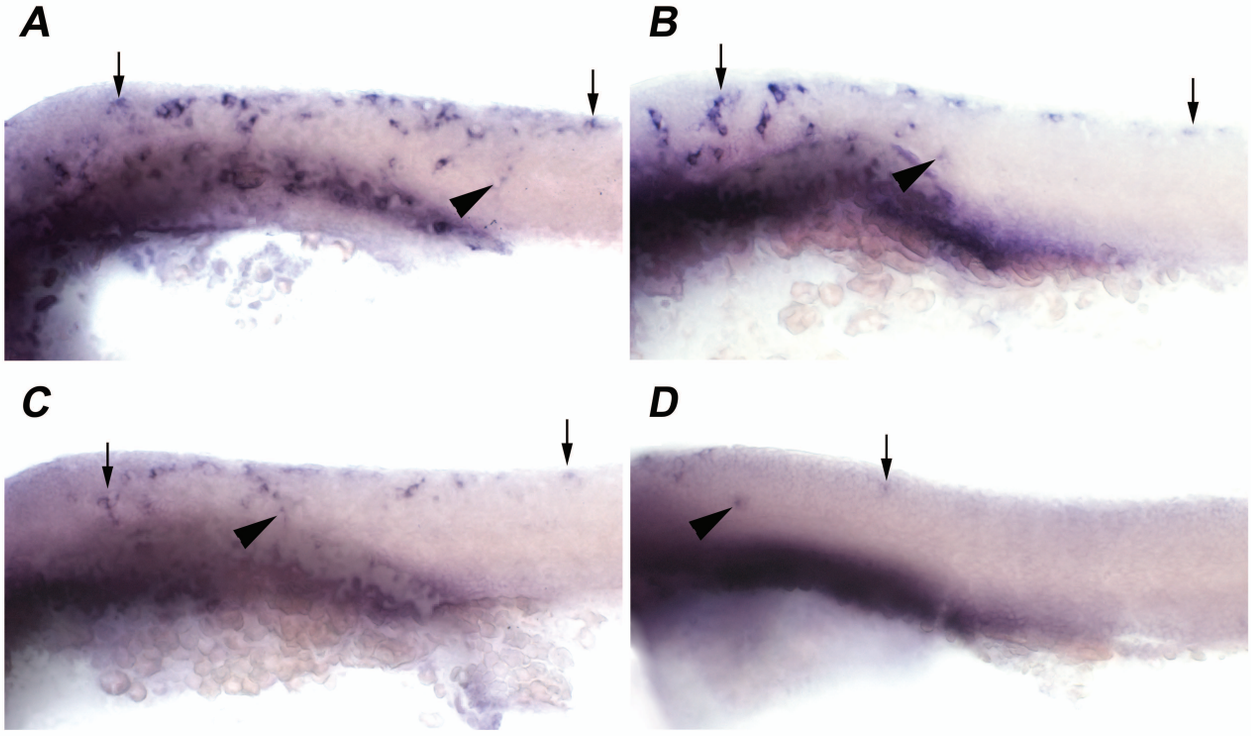


Figure 3.13. tfec upregulates pnp4a from 24 hpf. *In situ hybridisation results (anterior trunk region shown here) indicate variability in loss of pnp4a expressing cells in embryos injected with tfec MO. Arrows indicate the extent of the pnp4a positive NC domain along the dorsal trunk. Arrowheads indicate the extent of NC derivative migration. (A) shows the mildest observed phenotype amongst tfec morphants. pnp4a is expressed in several cells of the dorsal trunk (arrows) and in ventrally migrating cells (arrowhead). (B,C) indicate more strongly affected individuals. Dorsally located cells are labelled in both embryos (arrows), but migrating cells appear restricted towards the anterior (arrowheads). (D) shows the most severe phenotype, where a single weakly stained cell is present on the dorsal trunk (arrow) and another faint cell in detected along migratory pathways (arrowhead). Lateral views, oriented with the head to the left. Scale bar: 100 µm.*

3.3. Describing the preliminary GRN for iridophore specification

3.3.1. Mathematical modelling of intergenic relationships

The intergenic interactions described thus far were integrated to produce the preliminary iridophore GRN, as shown in Figure 3.14. Interactions within the network were mathematically described using a system of ODEs (refer to section 2.8.1).

Rate of change of *tfec* protein product over time

tfec was modelled to be cooperatively upregulated by Sox10 and by the formation of the Ltk/ligand complex. Although *tfec* and *mitfa* belong to the same gene family and Sox10 has been proven to bind to the *mitfa* promoter *in vivo* and upregulate *mitfa* expression (Potterf et al. 2000; Elworthy et al. 2003), it is not yet known whether Sox10 also binds the *tfec* promoter to directly upregulate *tfec* expression. Therefore, this interaction was represented using a dashed arrow. The Ltk/Ligand complex was shown to indirectly upregulate *tfec*, since intermediate signalling molecules and TFs are by definition required.

The following differential equation (1) was used to describe *tfec* expression dynamics:

$$\frac{d[T^P]}{dt} = g_T \times [r_S^T] \times (1 - [r_L^T]) + [r_L^T] - d_T \times [T^P] \qquad (1)$$

where $[T^P]$ represents Tfec protein concentration. The term g was used to designate the maximum rate of *tfec* expression, incorporating both mRNA and protein production. The rate of the latter was not explicitly modelled in this network. The term d refers to the protein degradation rate of Tfec.

The term $[r_S^T]$ was used to describe the effect of Sox10 on *tfec* expression as follows:

$$[r_S^T] = \frac{[S^P]}{Kd_S^T + [S^P]} \qquad (2)$$

and the term $[r_L^T]$ was used to describe the effect of the Ltk/ligand complex on *tfec* expression as follows:

$$[r_L^T] = \frac{[L^P]}{Kd_L^T + [L^P]} \quad (3)$$

$[S^P]$ represents the concentration of Sox10 protein and $[L^P]$ the concentration of Ltk/ligand complex.

In both (2) and (3), Kd values represent dissociation constants at equilibrium. These constants normally describe the affinity of a TF for a gene's promoter. For most interactions in this GRN, direct promoter binding has not been proven. Thus, these rates represented the efficiency with which the upstream effector (here Sox10 or Ltk/ligand complex) regulated the expression of its target gene (here *tfec*), even if this occurred through unknown intermediate interactions.

Rate of change of *ltk* protein product over time

Results from *in situ* hybridisation assays suggested that *ltk* is upregulated by Tfec. Whether Tfec directly binds to the promoter of *ltk* remains to be determined. Equation (4) was used to describe the dynamics of *ltk* upregulation:

$$\frac{d[Ltk^P]}{dt} = g_L \times [r_T^L] - d_L \times [Ltk^P] \quad (4)$$

where the term $[r_T^L]$ was used to describe the effect of Tfec on *ltk* expression as follows:

$$[r_T^L] = \frac{[T^P]}{Kd_T^L + [T^P]} \quad (5)$$

Ltk receptor/ Ltk ligand (L) binding and complex formation

The endogenous ligand of Ltk (denoted as 'L' in the GRN; Fig. 3.14) was assumed to form a complex with the receptor Ltk in order to indirectly activate transcription. The following expression describes the kinetics underlying receptor/ligand binding:

$$Ltk + Ligand \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} L^{complex} \quad (6)$$

In (6), k_{on} and k_{off} rates denoted the rates of association and dissociation between the two proteins, respectively. It followed that an appropriate equation to describe these dynamics was

$$\frac{d[L^{Complex}]}{dt} = k_{on} \times [Ltk] \times [Lig] - k_{off} \times [L^{Complex}] \qquad (7)$$

where the concentration of ligand ($[Lig]$) was assumed to remain constant, for simplicity and due to lack of more information at the time. The equation thus describing the rate of change of Ltk ligand concentration was

$$\frac{d[Lig]}{dt} = 0 \qquad (8)$$

Rate of change of *pnp4a* protein product over time

Experimental data suggested that both Tfec and Mitfa upregulate *pnp4a* expression. There was no experimental evidence to show that either TF binds the *pnp4a* promoter, thus both interactions were depicted using dashed arrows. The following differential equation described upregulation of *pnp4a* based on presence of either Tfec or Mitfa:

$$\frac{d[P^P]}{dt} = g_P \times ([r_M^P] + [r_T^P] \times (1 - [r_M^P])) - d_P \times [P^P] \qquad (9)$$

where $[P^P]$ represents the concentration of Pnp4a protein, the term $[r_M^P]$ was used to describe the effect of Mitfa ($[M^P]$) on *pnp4a* expression as follows:

$$[r_M^P] = \frac{[M^P]}{Kd_M^P + [M^P]} \qquad (10)$$

and the term $[r_T^P]$ was used to describe the effect of Tfec on *pnp4a* expression as follows:

$$[r_T^P] = \frac{[T^P]}{Kd_T^P + [T^P]} \qquad (11)$$

Rate of change of *mitfa* protein product over time

Sox10 has been previously shown to directly upregulate *mitfa* expression by binding to regulatory elements within the gene's promoter (Potterf et al. 2000; Elworthy et al. 2003). This interaction was, therefore, represented with a solid arrow in the network (Fig. 3.14) and *mitfa* expression dynamics were described using differential equation (12):

$$\frac{d[M^P]}{dt} = g_M \times [r_S^M] - d_M \times [M^P] \qquad (12)$$

where upregulation of *mitfa* by Sox10 was represented using the term

$$[r_S^M] = \frac{[S^P]}{Kd_S^M + [S^P]} \qquad (13)$$

3.3.2. Simulation of *sox10* expression dynamics

In this network, *sox10* was treated as an input and its expression dynamics were manually adjusted according to data from the extensive published work on the *sox10* expression pattern. Specifically, it has been shown that *sox10* is strongly expressed in premigratory NCCs at 18 hpf (Dutton et al. 2001) and that expression is gradually reduced throughout development of the melanophore lineage, a step which was proven to be crucial for differentiation of this pigment cell type (Greenhill et al. 2011). Furthermore, *sox10* expression had not been previously reported in iridophores, even though the expression pattern has been widely investigated (Greenhill et al. 2011; Takada and Appel 2011; Kwak et al. 2013). Therefore, at the first stages of this study, it was assumed that *sox10* is initially expressed in NC progenitors later giving rise to iridophores, gradually becoming downregulated as the lineage develops. This assumption is challenged in later stages (refer to section 4.4).

A Heaviside function was used to simulate high level of *sox10* expression during the early stages of iridoblast specification from the NC (18 hpf – 22 hpf), with the expression being gradually eliminated in an exponential manner following 22 hpf. In general terms, Heaviside functions allow for using different equations to describe how a quantity changes in distinct intervals.

Derivation of a generic Heaviside function

To derive a generic Heaviside function, a piecewise function (14) was first assumed

$$h(x) = \begin{cases} f(x) & for\ x \in [a, b) \\ g(x) & for\ x \in (b, c] \end{cases} \quad (14)$$

The following unit step function, $u(x)$, was used to re-define function $h(x)$:

$$u(x) = \begin{cases} 0 & for\ x < 0 \\ 1 & for\ x \geq 0 \end{cases} \quad (15)$$

First, the interval $[a, b)$ was considered. Upon consecutively substituting x with $x - a$ and $b - x$ in (15), equations (16) and (17) were obtained:

$$u(x - a) = \begin{cases} 0 & for\ x - a < 0 \Longrightarrow x < a \\ 1 & for\ x - a \geq 0 \Longrightarrow x \geq a \end{cases} \quad (16)$$

$$u(b - x) = \begin{cases} 0 & for\ x > b \\ 1 & for\ x \leq b \end{cases} \quad (17)$$

It follows that, for $a \leq x \leq b$

$$u(x - a) \times u(b - x) = 1 \quad (18)$$

whereas for $x < a$ or $x > b$

$$u(x - a) \times u(b - x) = 0 \quad (19)$$

Applying the same methodology for the interval $(b, c]$, the original piecewise function $h(x)$ (14) may be re-written in the form of a Heaviside function as follows:

$$h(x) = u(x - a) \times u(b - x) \times f(x) + u(x - b) \times (c - x) \times g(x) \quad (20)$$

Derivation of a Heaviside function to describe *sox10* expression dynamics

To generate a Heaviside function in order to simulate *sox10* dynamics, the generic $u(x)$ was replaced with the time-dependent function, $heaviside(t)$, that enabled definition of two time intervals in which *sox10* expression dynamics were described by two distinct equations $f(x)$ and $g(x)$. The intervals were chosen as follows:

1) 18 hpf (start of mathematical modelling, t=t_0) to 22 hpf (when a significant population of cells are likely to have begun becoming specified, gradually switching off premigratory NC markers)

2) 22 hpf to 100 hpf (differentiation stages should be reached, simulations are completed).

During the first interval, the concentration of Sox10 in the cells was simulated using the function

$$f(x) = g_s - d_s \times [S^P] \tag{21}$$

where g_s denotes the maximum rate of *sox10* expression, d_s the rate of Sox10 degradation and $[S^P]$ the average concentration of Sox10 protein in the cell population at a given time.

During the second interval, Sox10 concentration was made to exponentially decay by removing the production rate from the equation, thus obtaining:

$$g(x) = -d_s \times [S^P] \tag{22}$$

Hence, by substituting $u(x)$ for $heaviside(t)$, $[a, b)$ for $[18{,}22)$ and $(b, c]$ for $(22{,}100]$

the generic function $h(x)$ (14) generated the following differential equation which described *sox10* dynamics in the preliminary GRN:

$$\begin{aligned}\frac{d[S^P]}{dt} = {} & heaviside(t-18) \times heaviside(22-t) \times (g_s - d_s \times [S^P]) \\ & -heaviside(t-22) \times heaviside(100-t) \times d_S \times [S^P]\end{aligned} \tag{23}$$

3.3.3. Parameter choice

To numerically solve the system of ODEs, it was necessary to assign values for the various parameters featured in the equations. These were the following:

$\boldsymbol{Kd_A^G}$*:* The dissociation constant at equilibrium. By definition $Kd_A^G = b_A^G / a_A^G$, where

a_A^G*:* The rate of binding of A on the promoter of G.

b_A^G*:* The rate of unbinding of A from the promoter of G.

$\boldsymbol{g_G}$*:* The maximum rate of gene G expression. This rate incorporates mRNA and protein production levels, which were not explicitly modelled in this network.

$\boldsymbol{d_G}$*:* The degradation rate of protein encoded by gene G (G^P).

Since measuring the parameters biochemically was not feasible and estimating them using quantitative data posed several difficulties, a literature search for physiologically relevant values was carried out. Parameter values were derived from a number of different studies, if not relating to the genes of interest, then concerning different genes of the same family. The majority of the identified values were derived from *in vitro* experiments using mammalian proteins (tables 3.1 and 3.2), therefore they might not be accurate for describing the kinetics in the NC model system. Reassuringly, identified parameter values did not appear to deviate to a great extent, for example, DNA-protein interactions or protein-protein interactions maintained similar dissociation constants across studies and for different molecules (refer to tables 3.1 and 3.2).

3.3.3.1. Dissociation constants (Kd)

Literature-derived dissociation constants

1) Sox family genes.

It has been well established that Sox factors bind to DNA with low affinity (Kormish, Sinner, and Zorn 2010). During the literature search, estimated dissociation constants (K_d) of the Sox-4, Sox-5 and Sox-9 genes (van de Wetering et al. 1993; Connor et al. 1994; Mertin, McDowall, and Harley 1999) were identified (Table 3.1). The median of the three identified dissociation constants, thus a physiologically relevant parameter value to use for Sox family genes, was 2 nM.

2) MiT family genes

Mitfa and Tfec both belong to the MiT gene family and display highly conserved DNA-binding sequences (Hallsson et al. 2004; Pogenberg et al. 2012) and specificity for both E-box and M-box binding motifs (Martina et al. 2014). Only one relevant publication was identified (Pogenberg et al. 2012), in which binding affinities of MITF protein to its target

motifs were estimated (Table 3.1). The two identified constants, one for each target motif of MITF, were used to derive the median (2.1 nM), which was considered a physiologically relevant value to represent the affinity of MiT family genes for their targets.

3) RTK/Ligand interactions

The kinetics of RTK/ligand complex formation were investigated in an extensive body of work. To identify the k_{on} and k_{off} rates, the HEK (Kozlosky et al. 1995; Lackmann et al. 1996), MET (Dietz et al. 2013) and HTK (Bruno 1997) RTKs were considered (Table 3.2). Typical k_{on} rates for protein-protein interactions have been estimated between 10^6 – 10^7 M^{-1}x sec^{-1}. Therefore, using the literature-derived parameters and the relationship $K_d = k_{off}/k_{on}$, the required rates for RTK/Ligand interaction could be derived:

$$k_{on} = 10^{-2} \times 3{,}600\ nM^{-1} \times h^{-1}$$

$$k_{off} = 3 \times k_{on}\ h^{-1}$$

Table 3.1. Equilibrium dissociation rates for Sox and MiT family genes.

Gene Family	Gene	Species	K_d	Reference
Sox	Sox-4 Sox-5 SOX9	Mouse Mouse Human	0.03 nM 2 nM 12.4 ± 2.5 nM	(van de Wetering et al. 1993) (Connor et al. 1994) (Mertin, McDowall, and Harley 1999)
MiT	Mitf/M-box Mitf/E-box	Mouse	2.5 ± 0.5 nM 1.7 ± 0.2 nM	(Pogenberg et al. 2012)

Table 3.2. Affinity constants for RTK/Ligand interactions.

RTK	Ligand	Affinity constants	Reference
MET (HGFR)	HGF/SF	5 ± 0.8 nM (K_d)	(Dietz et al. 2013)
HEK	AL-1	20 nM (K_{on})	(Kozlosky et al. 1995)
HEK	AL-1	3.1 x 10^{-3} ± 0.6 sec^{-1} (K_{off}) 2-3 nM (K_d)	(Lackmann et al. 1996; Bruno 1997)
HTK	HTKL	c. 1 nM (K_d)	(Bruno 1997)

Adjusted dissociation constants

Although the literature-derived constants were considered biologically relevant, these values had to be modified in order for the mathematical model to be able to simulate gene expression dynamics. The system of ODEs describing model A was implemented in MATLAB software using the literature-derived K_d values and the rates described in the following section. As a result, gene expression failed to become activated and, instead, all protein concentrations rapidly reached a plateau stage at 0 nM (data not shown). Different tests revealed that, given the chosen parameter values for maximum gene expression and protein degradation rates (refer to next section), in order to achieve gene activation the literature-derived dissociation constants had to be decreased by one order of magnitude (data not shown). Therefore, K_d values were decreased from 2 nM to 0.2 nM, implying higher binding affinity of a TF, which translates to more efficient gene activation.

It should also be noted that efficiency of Sox10-dependent activation was assumed to be somewhat higher than the strength of MiT factor binding, in order to account for cooperative interactions taking place *in vivo*, which have not been incorporated in the model. Specifically, Sox10 has been reported to dimerise on promoters of target genes (Schlierf et al. 2002; Ludwig, Rehberg, and Wegner 2004) and to interact with co-factors such as Pax3 in order to enhance the binding affinity of the transcriptional complex (Potterf et al. 2000; Bondurand et al. 2000). Therefore, the Sox10 dissociation constant was assumed to be 10% higher than that of MiT factors Mitfa and Tfec.

Finally, the efficiency with which the Ltk/ligand complex is able to indirectly activate transcription is not known. For simplicity, the respective rate was assumed to be equal to the one assigned to MiT factors.

Overall, the following parameter constants, which were chosen to a certain extent arbitrarily but still closely resembled physiologically relevant values according to the literature, were used in the model:

Sox10 binding to promoters of downstream factors: $K_d{}^S = 0.2\ nM$.

Tfec binding to promoters of downstream factors: $K_d{}^T = 0.3\ nM$.

Ltk/Ligand complex indirectly upregulates downstream factors: $K_d{}^L = 0.3\ nM$.

Mitfa binding to promoters of downstream factors: $K_d{}^M = 0.3\ nM$.

3.3.3.2. Maximum expression rate (g) and degradation rates (d)

Maximum gene expression rate (g)

Global gene expression analysis (Schwanhäusser et al. 2011) suggested that cellular transcription rates demonstrated a median of 2 mRNA molecules produced per hour. For different genes, this rate was found to span two orders of magnitude. To use this rate, conversion of the units to nM/hour was required in order to maintain consistency with other units.

Using FIJI, DIC images (20x magnification) of embryos in which nuclei were labelled with DAPI were analysed and the average cytoplasmic volume of cells residing in NCC positions along the dorsal trunk was measured at the following stages:

- 18 hpf (n=5 cells; mean cytoplasmic volume 186.97 μm^3)
- 24 hpf (n=8 cells; mean cytoplasmic volume 257.41 μm^3)
- 36 hpf (n=7 cells; mean cytoplasmic volume 126.19 μm^3)
- 4 dpf (n=6 cells; mean cytoplasmic volume 73.88 μm^3)

Cell width and length followed by nuclear width and length were measured using FIJI tools. The cells were assumed to be shaped as rods with rounded ends (based on DIC Z-stack visualisation, data not shown), and therefore the following formula was used to calculate cytoplasmic volumes:

$$Volume(\mu m^3) = \pi \times W^2 \times \frac{L - (W/3)}{4} \qquad (24)$$

where W stands for cell width and L stands for cell length. The nuclear volume was subtracted from the total cell volume, providing an estimate of the cytoplasmic volume.

Across stages, the average cytoplasmic volume was 161.11 μm^3, or 0.16 pL. Based on the 2 molecules/hour rate (Schwanhäusser et al. 2011), the average gene expression rate of 0.02 nM/hour was thus derived (1 molecule = 1.7 x 10^{-15} nmoles, hence 2 molecules/hour = 3.4 x 10^{-15} nmoles/hour therefore by dividing with the average cytoplasmic volume ≈0.02 nM/hour was obtained). The maximum gene expression rate was arbitrarily assumed to be 10 times higher than the calculated average rate. For all genes the following maximum production rate was thus assumed:

$$g_G = 0.2\ \frac{nM}{h}$$

Protein degradation rates (d)

Degradation rates were estimated based on protein half-lives reported in published literature (Mertin, McDowall, and Harley 1999; Belle et al. 2006; Eden et al. 2011; Schwanhäusser et al. 2011). The calculated half-life median in one study was 46 h, whereas the smallest measured half-life was 0.43 h (Schwanhäusser et al. 2011). Others found that protein half-lives followed a log-normal distribution with mean and median of approximately 0.72 h (Belle et al. 2006), or that they could vary between 0.75-22.5 h (Eden et al. 2011). Moreover, the half-life of human SOX9 protein in cell lines was reported to be 3.6 ± 0.22 h (Mertin, McDowall, and Harley 1999). To convert half-lives to degradation rates (h^{-1}), published guidelines were followed (Kuhar 2010), acco (25) which:

$$d = 0.693/t$$

where d was the degradation rate constant and t was the experimentally estimated half-life of a protein of interest. Table 3.3 displays the values generated using formula (25) to convert the published half-lives into degradation rate constants.

Table 3.3. Degradation rate constants.

	Mertin et al. 1999	**Schwanhäusser et al. 2011**	**Eden et al. 2011**	**Belle et al. 2006**
d_{min}	0.1814 h^{-1}	0.0151 h^{-1}	0.0308 h^{-1}	-
d_{max}	0.2050 h^{-1}	1.6116 h^{-1}	0.9240 h^{-1}	-
d_{mean}	0.1932 h^{-1}	0.8133 h^{-1}	0.4774 h^{-1}	0.9625 h^{-1}
overall d_{mean}	0.6116 h^{-1}			

Taking into account the average calculated d_{mean} (0.6116 h^{-1}) and the fact that the degradation rate for SOX9, a TF relevant to this study, was estimated at approximately 0.1932 h^{-1} (Mertin, McDowall, and Harley 1999), values between 0.1 and 0.6 h^{-1} were arbitrarily selected as likely physiologically relevant degradation rate constants. Furthermore, considering that TFs are inherently much less stable than other proteins, such as enzymes and transmembrane receptors (Belle et al. 2006), they were assigned higher degradation rates.

For the mathematical modelling, the following degradation rates were selected for each gene of interest:

Sox10 degradation rate: $d_S = 0.5\ h^{-1}$

Tfec degradation rate: $d_T = 0.5\ h^{-1}$

Ltk receptor degradation rate: $d_L = 0.25\ h^{-1}$

Mitfa degradation rate: $d_M = 0.5\ h^{-1}$

Pnp4a degradation rate: $d_P = 0.1\ h^{-1}$

3.3.3.3. Initial values

It was then necessary to determine the initial values for t=t_0, hence the average concentrations of each protein coded by each of the genes of interest at 18 hpf. Owing to the lack of an appropriate method to quantify such values, *in situ* hybridisation experiments were used to determine whether expression of each gene was present or absent at that stage at 18 hpf (Fig. 3.15). Because development occurs gradually from anterior to posterior, a region of the posterior dorsal trunk (Fig. 3.15 B,D,F,H,J) was arbitrarily specified to ensure a more developmentally homogeneous population of premigratory NCCs was considered to begin the modelling.

Strong *sox10* and *tfec* expression was detected in the selected region of interest. The two genes were expressed in the majority of cNCCs and tNCCs at this stage (Fig. 3.15 A-D). *ltk* expression was only detectable in very few cNCCs and in association with the presumptive RPE cells at 18 hpf, as well as in the notochord (Fig. 3.15 E,F). *mitfa* expression was detectable in the cNC and presumptive RPE domains and in a few emerging tNCCs towards the anterior trunk, but was not yet observed in the selected region (Fig. 3.15 G,H). Finally, *pnp4a* expression was only seen in very few cells of the head, but not in the posterior dorsal trunk region (Fig. 3.15 I,J).

Therefore, the following initial values were determined for the mathematical model:

Sox10 (gene strongly expressed in specified region): $S_0 = 0.4\ nM$

Tfec (gene expressed in specified region, but likely less strongly than *sox10* based on *in situ* hybridisation signal intensity): $T_0 = 0.25\ nM$

Ltk (gene expression undetectable in specified region): $Ltk_0 = 0\ nM$

Ltk ligand (uncharacterised): $Ligand_0 = 1\ nM$

Ltk/ligand complex: $L(complex)_0 = 0\ nM$

Mitfa (gene expression undetectable in specified region): $M_0 = 0\ nM$

Pnp4a (gene expression undetectable in specified region): $P_0 = 0\ nM$

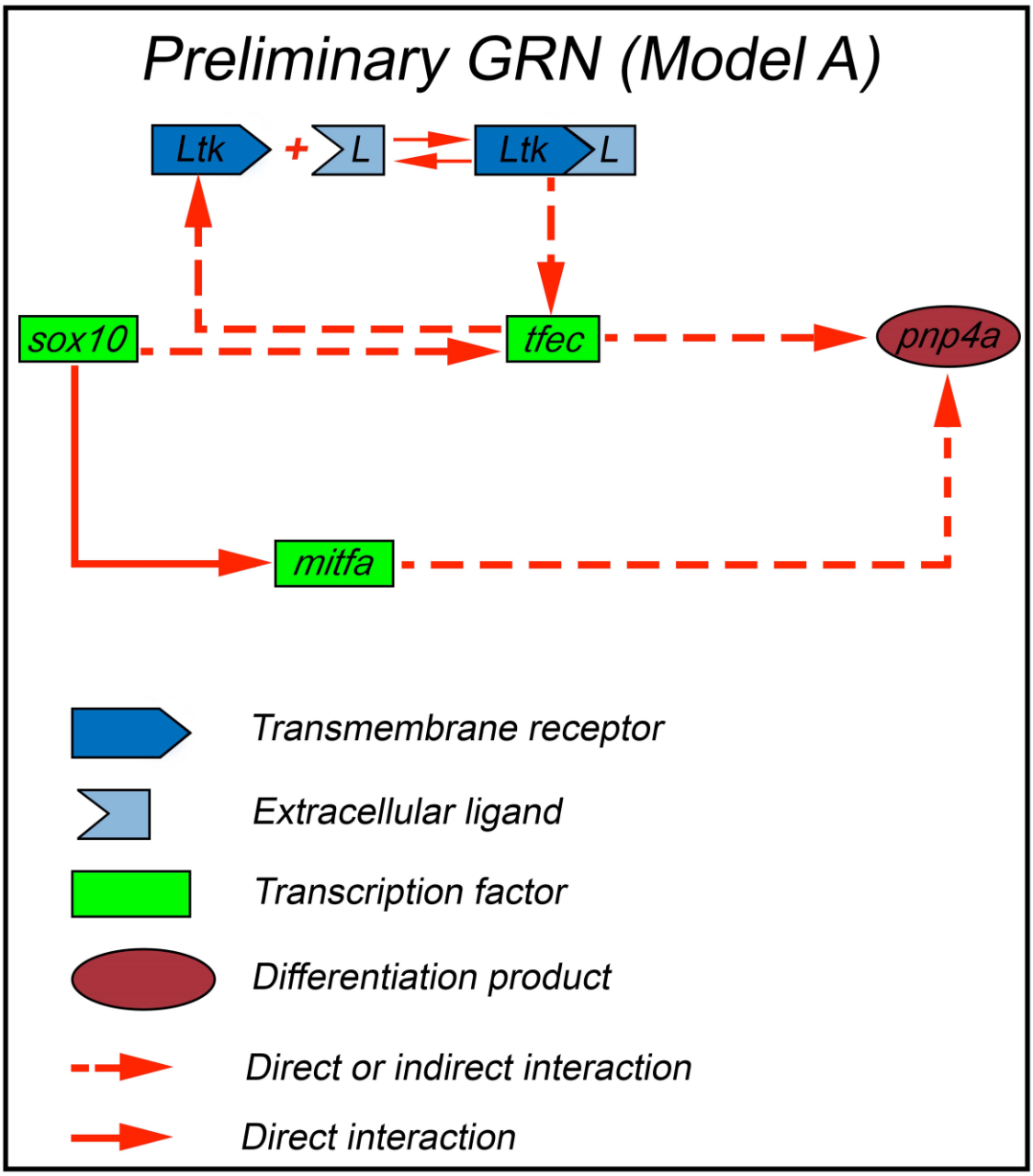


Figure 3.14. The preliminary iridophore GRN. The key indicates symbols used for different types of molecules and interactions. All nodes are represented using gene names (lower case, italised), except for Ltk, which is shown as the RTK protein (capital first letter) binding to its ligand (L) to form a complex that activates gene expression. At equilibrium, Ltk and ligand form the complex and dissociate at specified rates (k_{on}, k_{off} respectively).

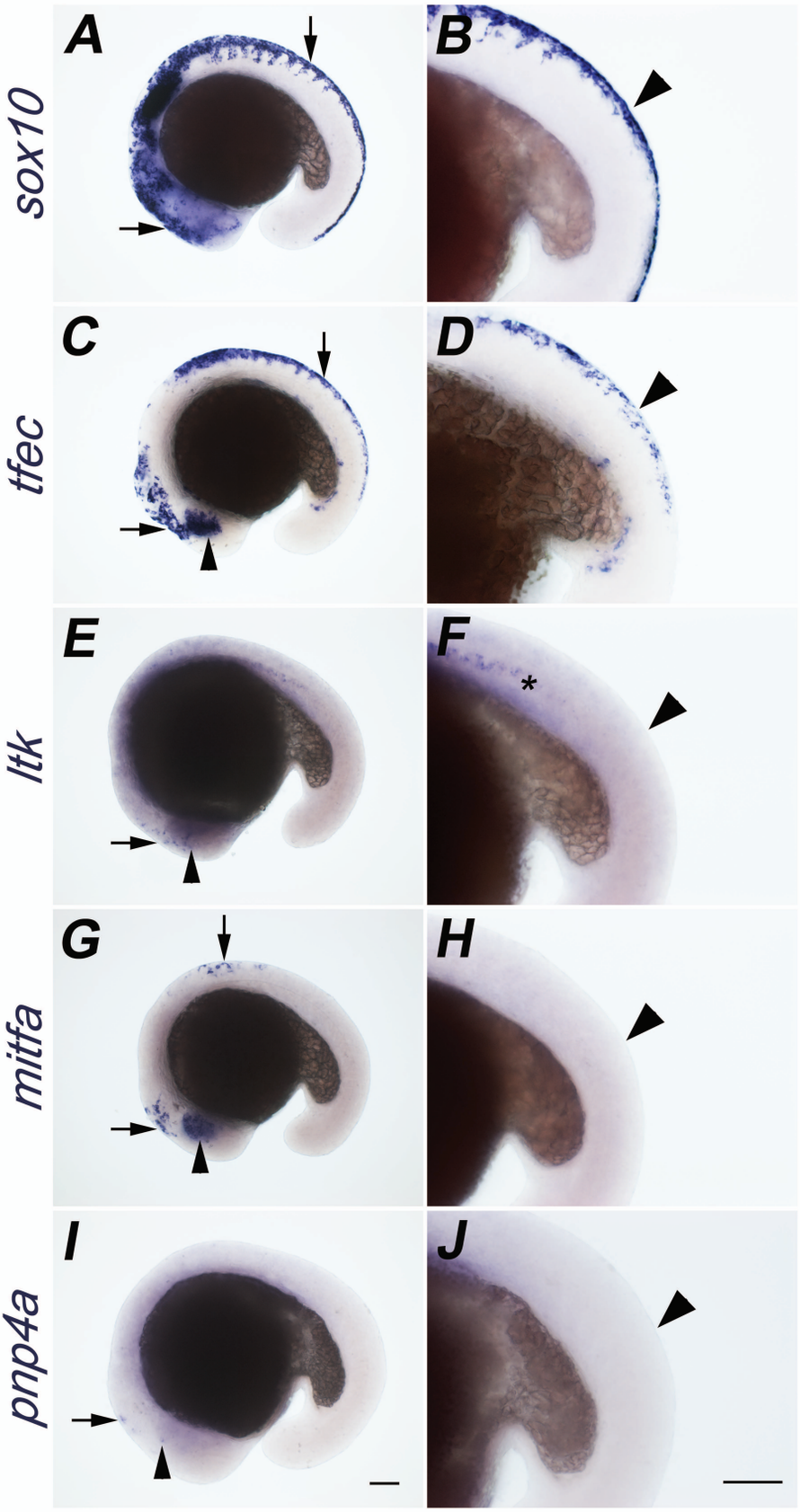


Figure 3.15. sox10 and tfec, but not ltk, mitfa and pnp4a are expressed at 18 hpf. *In situ hybridisation results show expression of candidate genes of interest at t0 (18 hpf). In (A, C, E, G, I), stained cNCCs and tNCCs are indicated by arrows and the presumptive RPE region by vertical arrowheads. In B, D, F, H, J, arrowheads indicate the region specified to determine initial expression values. In (F) asterisk denotes expression of ltk in the notochord. Lateral views, oriented with the head to the left. Scale bars: 100 μm.*

The initial concentrations values for Sox10 and for Tfec were chosen as indicated above on the basis of preliminary testing of the mathematical model. Briefly, the selected values produced smooth curves (data not shown), without causing the initial concentration of each gene to decline or rise sharply in the beginning of the simulations. Importantly, choosing alternative initial values did not result in alternative plateau phases (data not shown), thus the output of the model was not affected by this choice of parameters. It should be noted that the concentration of the currently unknown ligand of Ltk was set to 1 nM. For simplicity and to avoid making unjustified assumptions, the ligand concentration was assumed to remain constant and abundant in this model.

3.3.4. Computational implementation

To numerically solve the equations and to simulate the predicted gene expression dynamics over time in the iridophore lineage, MATLAB software was employed (refer to section 2.8.2). To simulate gene A loss of function, $\left[r_A^G\right] = 0$ was set to indicate that it is no longer able to elicit effects on the promoters of a downstream target G. The predicted dynamics for model A are shown in Figure 3.16. When interpreting these simulations, it should be noted that neither absolute concentration levels (y-axis), nor times (x-axis) were considered to be accurate quantitative representations, since parameter values were arbitrarily chosen and quantitative data were unavailable. Therefore, only the shape of the curves was considered at this stage.

In a WT context, genes expressed in differentiated iridophores, namely *tfec*, *ltk* and *pnp4a* (refer to sections 3.2.1, 3.2.3 and 3.2.6, respectively), were successfully predicted to remain upregulated throughout development of the lineage (Fig. 3.16; Wild type). Furthermore, Sox10 concentration was set to exponentially decrease following 22 hpf (refer to section 3.3.2), followed by a corresponding decrease in Mitfa. This prediction was considered biologically relevant, because *mitfa* is the melanophore master regulator and its expression was not expected in the iridophore lineage (J A Lister et al. 1999).

When loss of *sox10* function was simulated (Fig. 3.16; *sox10*$^{-/-}$), *mitfa* was shown to never become upregulated. This prediction was consistent with biological observations according to published literature (Dutton et al. 2001). All iridophore-specific genes (*tfec, ltk* and *pnp4a*) were maintained throughout development and Sox10 protein levels remained unaffected. Indeed, Dutton *et al*. have previously shown that an increased number of NCCs express *sox10* transcript in *sox10* mutants (Dutton et al. 2001). Moreover, experiments presented in sections 3.2.2 and 3.2.3 revealed that both *tfec* and *ltk* remained activated for a prolonged period of time in *sox10* mutants in unspecified NCC progenitors. Thus, predictions of model A in the *sox10* mutant context were

consistent with data used to build model A, as well as with results from the published literature to date.

Loss of *ltk* function resulted in an initial rise of Tfec and Pnp4a, but also of Ltk concentrations, which declined after approximately 30 hpf. Data presented in this chapter supported the predicted behaviour of *tfec* (section 3.2.5), while published work rendered the predicted *ltk* dynamics biologically relevant (Lopes et al. 2008). Next, loss of *tfec* function was simulated, which demonstrated a similar predicted phenotype. In this context, Pnp4a and Tfec concentrations gradually declined while *ltk* was never expressed. Although morpholino-mediated knockdown experiments suggested that the predictions for Ltk and Pnp4a dynamics were likely to be valid (sections 3.2.5 and 3.2.6), *tfec* mutants are required to thoroughly evaluate the predictions of model A in this context. Finally, loss of *mitfa* function was predicted to have no significant effect on iridophore genes, all of which started being produced without significant delay and reached their respective WT plateau levels. Published work has demonstrated that morphologically normal iridophores are present in *mitfa* mutant embryos (J A Lister et al. 1999), thus the predictions were once again considered consistent with experimental observations.

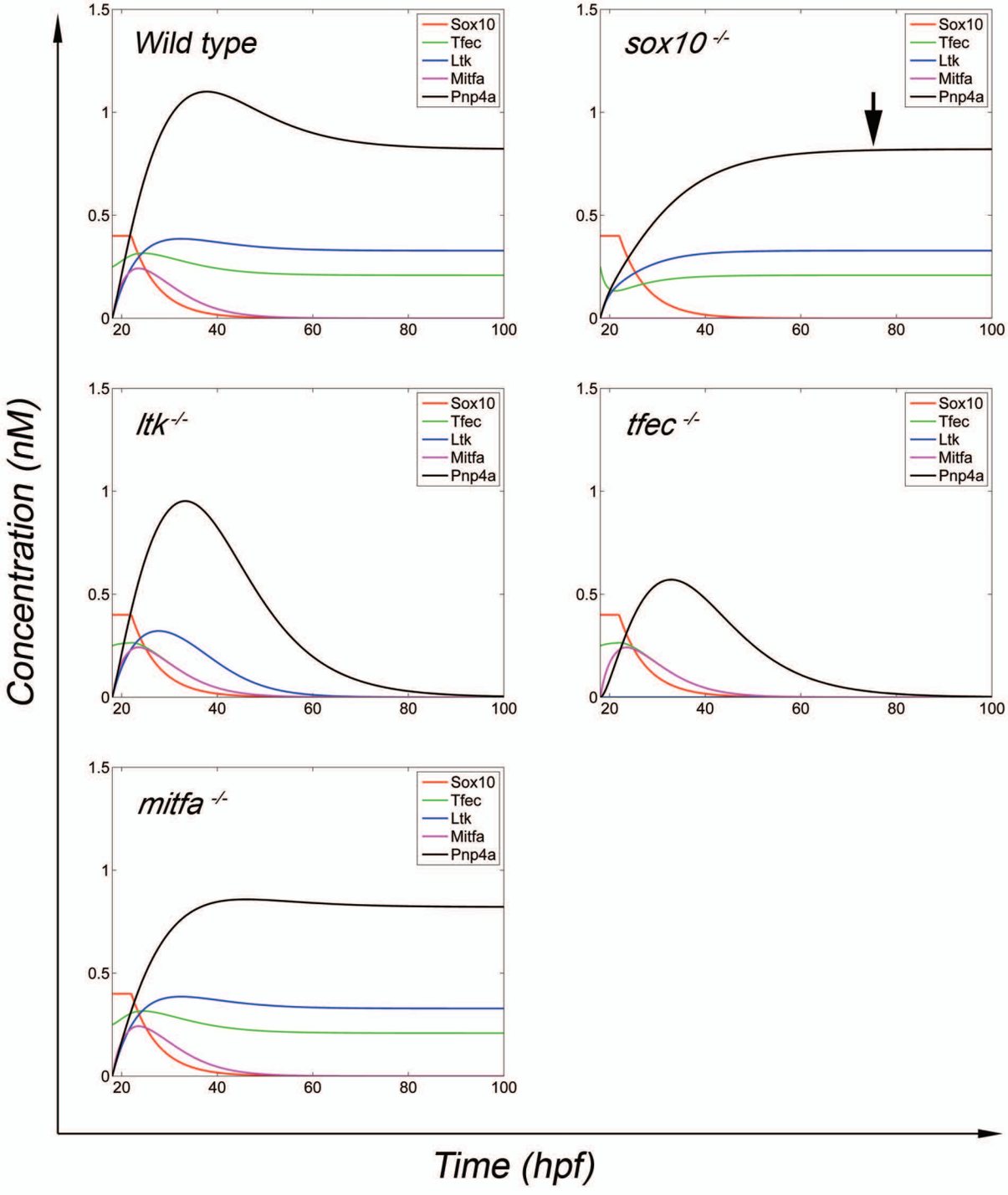


Figure 3.16. Predictions of gene expression dynamics in model A. *The diagrams show how protein concentrations (nM; y-axis) change over time (hpf; x-axis) in the model for WT and loss of gene function predictions, as indicated in the top left of each diagram. Different genes are represented by individual colours as shown in the legend (top right of each diagram).*

3.4. Discussion

In this chapter, a set of candidate genes of interest were defined and interactions between them were identified in order to build a preliminary GRN guiding iridophore development. Using primarily *in situ* hybridisation on WT zebrafish embryos, as well as embryos presenting with loss of function of candidate genes of interest, a *sox10*-dependent positive feedback loop between *ltk* and *tfec*, which was suggested to be crucial for iridoblast survival and differentiation, was established. Experimental evidence led to generating Model A, which was mathematically described using a system of ODEs and simulated using MATLAB software. Computationally simulating model A generated predictions regarding how gene expression dynamics progress over time in the model. These predictions were subsequently compared to the patterns of gene expression published in the literature and determined experimentally over the course of this project, in order to evaluate their consistency with known biology.

3.4.1. *sox10* upregulates *tfec* in the iridoblast

In keeping with what was previously reported in the literature (James A Lister et al. 2011), *tfec* expression at 18 hpf was detected along the premigratory cNC and tNC domains. From 24 hpf, cells which had likely become fate-restricted towards the iridophore lineage maintained *tfec* expression and could be detected migrating towards the ventral side of the trunk. Concomitantly, *tfec* expression was downregulated in cells becoming specified towards other fates. At this stage, presumptive iridoblasts prominently expressing *tfec* were identified strictly along the medial pathway (Robert N Kelsh 2004). *tfec* expression remained weakly detectable in cells migrating through the lateral pathway, and along the dorsal premigratory domain, indicating that *tfec* was gradually becoming downregulated in other NC derivatives as specification progressed. At following time points *tfec* expression was found in presumptive specified iridoblast positions and, ultimately, labelled mature iridophores (Lister et al. 2011; this work).

In *sox10* mutants, *tfec* expression failed to be upregulated in specified, migrating iridoblasts. In addition, overexpression data by Dr T. Subkhankulova suggested that *sox10* is able to upregulate *tfec* expression. Thus, it was concluded that *tfec* maintenance in iridoblasts is *sox10*-dependent. Interestingly, this interaction only became apparent in the context of the specified iridoblast. In *sox10* mutants at 18 hpf, expression of *tfec* across the premigratory NC domain was unaffected, and failed to be repressed as specification progressed. Similarly, *foxd3*, *sox9b* and *tfap2a* single mutants did not display alterations in *tfec* expression within premigratory NCCs at 18 hpf (this work, data not shown). An explanation is that *tfec* expression in early NCCs is redundantly activated by factors functioning during NC establishment. To definitively test this, investigation of double mutants would provide useful evidence.

To that end, the *sox10*^{t3};*sox9b*^{fh313} double mutant line was generated over the course of this project and awaits examination. Interest for this line comes from the fact that redundant functions have been previously demonstrated between *sox9b* and *sox10* both in neuronal and in melanophore development (Carney et al. 2006; Greenhill et al. 2011), while both genes are members of the NC specification GRN (refer to section 1.3.2). Therefore, it is likely that functional redundancy is also present in the context of the NC. Furthermore, f*oxd3*;*tfap2a* double mutants have been previously characterised and functional redundancy between *foxd3* and *tfap2a* has been demonstrated during NC specification (Arduini, Bosse, and Henion 2009). Finally, in *Xenopus* embryos, Pax3 and Zic1 have been shown to cooperate to upregulate a host of NC specifier genes (Sato, Sasai, and Sasai 2005; Plouhinec et al. 2014). It would be instructive to examine *tfec* expression in these loss of function contexts to determine whether the same pairs of genes act redundantly to activate *tfec* in multipotent NCCs. It may, however, be the case that elimination of *tfec* expression is still not observed when the function of two genes is simultaneously ablated. Redundancy between three or more transcriptional regulators could be tested by using appropriate combinations of morpholinos, or the CRISPR/Cas9 technology, to abolish the function of genes of interest and by assessing subsequent changes in the early expression of *tfec*.

Strikingly, the *tfec* expression pattern in *sox10* mutant embryos was found to be reminiscent of the previously established *ltk* pattern in this genetic background (Lopes et al. 2008). *ltk* has been proposed to function in a presumptive common progenitor, called the chromatoglioblast (Lopes et al. 2008; Nikaido et al., in prep.), which in *sox10* mutants becomes 'trapped' in the premigratory NC domain upon failure to become specified (Lopes et al. 2008; this study). Therefore, it was considered that *tfec* might label the same common progenitor. Although there is convincing evidence in support of the chromatoglioblast (Nikaido et al., in prep.), definitive confirmation of its existence and of expression of either *ltk* or *tfec* during its development requires generation of appropriate transgenic lines, allowing for permanently labelling all derivatives of cells that transiently expressed either *ltk* or *tfec*. Similar transgenic lines, labelling all derivatives of *sox10* positive NCCs, have been established and widely used (Rodrigues et al. 2012; Mongera et al. 2013).

3.4.2. *ltk* and *tfec* regulate each other via a positive feedback loop

A positive feedback loop between *ltk* and *tfec* expression was identified. Loss of either gene's function resulted in elimination of the other gene's expression. More specifically, in *ltk* mutants, initial expression of *tfec* in the NC was unaffected and migrating iridoblasts were detectable at 24 hpf. This result suggested that specification had occurred. However, gradual reduction in the number of *tfec* positive cells followed, indicating that

survival of specified cells was impaired. This *sox10*-dependent *tfec*/*ltk* positive feedback loop was at the core of the preliminary GRN and closely resembled structural features of the melanophore GRN. In melanophores, a positive feedback loop triggered by *sox10* between *mitfa* and factor Y, which was later suggested to correspond to the RTK *kit*, was identified as crucial for melanoblast specification and survival (Greenhill et al. 2011; A. Lapedriza, PhD thesis). To further investigate this feedback loop, generation of *tfec* mutants is required. The non-specific effects and the high degree of variability inherent to morpholino-mediated knockdown prevent rigorous assessment of subtle effects. For example expression in a small subset of cells might be overlooked or obscured. Generation of *tfec* mutants is described in the next chapter.

When considering the *ltk*/*tfec* feedback loop, questions arise as to why in WT embryos *tfec* does not activate *ltk* expression in the premigratory NC. One possible explanation is that, at this stage of NC development, a repressor prevents upregulation of *ltk*. This role could be assumed by a TF functioning in the NC establishment GRN (refer to section 1.3.2), responsible for preventing precocious fate restriction. Alternatively, it is likely that, in order to activate *ltk* expression, *tfec* requires a cofactor, which in not present in early NCCs. One method to distinguish between these two possibilities, would be by overexpressing *tfec* in the early embryo and testing whether it is able to upregulate *ltk* in a non-NC context. If this is the case, the existence of a NC-specific repressor would be supported. If *tfec* is unable to upregulate *ltk* in a cell type different than the iridoblast, it is plausible that an essential cofactor has been overlooked. Since no good candidates have been identified for this role, conducting either RNA-Seq in a purified population of iridoblasts or iridophores, or appropriate ChIP-seq experiments might prove instructive to highlight previously unknown candidates for the GRN.

Finally, it should not be disregarded that functionality of *tfec* in the premigratory NC has not been demonstrated thus far. Therefore, a third explanation for the inability of *tfec* to trigger *ltk* in the premigratory NC might be that *tfec* function is inhibited at the protein level. Intriguingly, expression of the bHLH inhibitors *id2a* or *id3* has been demonstrated in the premigratory NC domain (F. Rodrigues, PhD thesis). These Id (inhibitor of differentiation) family TFs might interact with Tfec directly, or with factors downstream of or interacting with Tfec, crucial for the TF to elicit its transcriptional effects. As a result, Id proteins would inhibit *tfec* function and prevent precocious iridoblast development. To further investigate this hypothesis, loss of function of *id2a* and *id3* must be induced, followed by assessment of *ltk* expression in the premigratory NC domain. A useful complementary experiment would be to determine whether these factors are capable of physically interacting with *tfec*, for example by performing affinity precipitation of protein complexes.

Interestingly, upon loss of *sox10* function, prominent *ltk* expression is detected in the premigratory NC domain at 24 hpf, in a manner reminiscent of *tfec* expression (Lopes et al. 2008; this work). It may be the case that, in *sox10* mutants, the repression of *tfec*-mediated *ltk* expression in NCCs is abolished and the *tfec*/*ltk* positive feedback loop is triggered, thus facilitating the switch from an early NCC expressing *tfec* to a partially restricted progenitor capable of generating iridophores. Alternatively, *ltk* expression might be independent of *tfec*. To determine whether *ltk* and *tfec* depend on each other for their activation in this context, examining the *tfec* and *ltk* expression phenotypes in *sox10*;*ltk* and *sox10*;*tfec* double mutants, respectively, is imperative. If establishment of the *tfec*/*ltk* feedback loop in trapped NC-derived progenitors in *sox10* mutants is proven, this provides an explanation of why *ltk* expression is prominent in this cell population, even though its expression has not been detected in premigratory NCCs, with the exception of a relatively small number of cells at 24 hpf (Lopes et al. 2008).

3.4.3. *pnp4a* activation is both *tfec* and *mitfa*-dependent

It was shown that both the melanocyte master regulator, *mitfa*, and the hypothesised iridophore master regulator, *tfec*, are important for the upregulation of *pnp4a,* starting from early NCC specification stages. *tfec*-dependent upregulation of *pnp4a* was not surprising, since *pnp4a* was reportedly an early marker of the iridophore lineage (Curran et al. 2010). However, for this reason, *mitfa*-dependent upregulation was unexpected. Interestingly, data presented in this chapter indicated that *mitfa* is only able to upregulate *pnp4a* during early stages of specification of NC derivatives, while *tfec*-mediated *pnp4a* activation is sustained, resulting in *pnp4a* expression readily labelling differentiated iridophores. These temporally restricted effects on gene regulation suggest that *pnp4a* activation might require additional cofactors. This notion is further investigated and discussed in more detail in chapter 4.

3.4.4. Advantages and limitations

3.4.4.1. Experimental methodology

For the purposes of constructing the preliminary GRN governing iridophore development, chromogenic *in situ* hybridisation was the technique of choice. Several advantages are associated with this experimental methodology, as well as important limitations.

Advantages

Importantly, *in situ* hybridisation allows for spatio-temporal characterisation of gene expression. Specifically, types of cells expressing a gene of interest can be identified by their exact position within the embryo, which is particularly instructive in a highly

migratory cell type like the NCC, as well as their morphology in certain cases. For example, NCCs and NC derivatives can be distinguished based on their migratory route of choice and their respective positions along the antero-posterior and the dorso-ventral axes. This facilitates characterisation of mutant phenotypes manifesting even in small subsets of cells expressing a gene, which may not have been obvious if the embryo was dissociated. To illustrate this point, although the elimination of *mitfa* expression in *sox10* mutants (Elworthy et al. 2003) could be efficiently determined by performing q-RT PCR in WT and mutant embryos, this would not be the case for *tfec* in the same mutants. Although upon loss of *sox10* function *tfec* expression is abolished in the relatively few specified iridoblasts which are normally present in the embryo, it is maintained in a significant number of trapped precursors along the premigratory NC domain. Therefore, it would be misleading to examine the effects on *tfec* activation without considering the position of cells expressing the gene. In other cases, a gene of interest may be functioning, yet differentially regulated, in different organs. For instance, *tfec* is active in the ICM, although none of the NC-related mutant phenotypes affect its expression in this tissue (Lister et al. 2011; this work). Moreover, *ltk* transcript is detectable in the notochord prior to onset of expression in the NC (Lopes et al. 2008; this work). Finally, expression of *id2* has been demonstrated in a variety of different tissues other than the NC (Chong et al. 2005; F. Rodrigues, PhD thesis).

Importantly, cells within each anterior-posterior region of the embryo are of a different developmental stage, with anteriorly located cells being more developmentally advanced than cells towards the posterior. Therefore, techniques involving dissociation of the entire embryo can prove dangerous due to averaging populations of more or less progressed cells at different stages of development. This is a matter of particular concern when investigating gene expression changes during early developmental stages, when cells undergo dynamic changes. Using *in situ* hybridisation, WT and mutant phenotypes can be interpreted taking into account the stage at which a certain phenotype manifests.

To a certain extent, these effects could be bypassed by dissecting parts of the embryo for quantitative analyses (A. Lapedriza, PhD thesis), although this is not possible in all stages and can often lead to experimental artefacts stemming from slight variations in manipulation and handling between different specimens. Also, considering the example of *tfec* expression in *sox10* mutants, dissecting even very small parts of the trunk would not produce a definitive result by q-RT PCR.

Disadvantages

Despite the aforementioned benefits, using *in situ* hybridisation to identify intergenic interactions presents with important drawbacks. Firstly, the sensitivity threshold for detecting gene expression is often high in relation to that of quantitative methods, such

as q-RT PCR. As a result, low-level gene expression in a particular population might be overlooked. In this study, extensive optimisations of experimental conditions were carried out, with the aim of lowering the sensitivity threshold. Consequently, novel features of expression patterns could be discerned. For example, putative downregulation of *tfec* expression along the premigratory NC domain at 24 hpf, as well as the presence of laterally migrating cells with low levels of *tfec* expression could be recognised. Importantly, *pnp4a* was previously considered an exclusive marker of the iridophore lineage based on its expression pattern (Curran et al. 2010). However, optimisation of the experimental conditions led to identifying expression of the gene in migrating melanoblasts, as the pattern was strongly reminiscent of *mitfa* expression at early stages. Despite these improvements, it remains likely that certain aspects of gene expression remain unknown. In chapter 4, RNAscope, a newly established technique for use with whole-mount zebrafish embryos, which was reportedly associated with significantly lowered sensitivity thresholds, is introduced.

Another important disadvantage of *in situ* hybridisation is its inherent inability to yield quantitative data. This is because the chromogenic compound employed for staining diffuses within the cells and fills the cytoplasm, while its production does not strictly correspond to the abundance of transcript within cells. Therefore, qualitative data could only be produced and the outputs of the model could not be assessed in terms of predicted concentrations of genes of interest. Due to the lowering of the sensitivity threshold in this study, which led to staining of some cells appearing weaker than in others, semi-quantitative predictions could, at best, be tested.

Finally, a major drawback of the experimental methodology used here was the reliance upon presence of mRNA to formulate hypotheses regarding the function of TFs. It should not be disregarded that the rate of translation and of mRNA and protein turnover, as well as post-translational modifications likely impact on the functionality of the protein product of an expressed gene. Reassuringly, it was recently indicated that the abundances of transcripts and corresponding proteins are clearly correlated (Schwanhäusser et al. 2011). To better be able to characterise the functional role of gene expression within GRNs characterisation of protein abundance is necessary. This can be carried out using immunohistochemistry, although in the zebrafish model, appropriate antibodies for most genes have yet to become available.

Overall, for the purposes of this project it was crucial to accurately determine the spatio-temporal expression patterns of genes of interest. However, the qualitative results obtained by *in situ* hybridisation did not allow for achieving a high degree of predictive accuracy when analysing computational simulations. Specifically, the biological accuracy of predicted concentration values could not be evaluated using the presented

experimental evidence. Moreover, lack of quantitative data prevented parameter estimation, which could be performed by fitting curves to the data (Fomekong-Nanfack et al. 2009).

3.4.4.2. Quantitation of gene expression in the iridophore lineage

To achieve quantitative results, absolute gene expression levels need to be investigated using q-RT PCR. This is not at the moment possible due to lack of appropriate resources allowing for isolation of cell populations of interest from the developing embryo. To overcome this limitation, transgenic lines must be generated, in which cells of the iridophore lineage are fluorescently labelled and can be isolated via fluorescence activated cell sorting (FACS). To that effect, genes which are expressed uniquely in the iridophore lineage must be identified. *tfec* is not an appropriate candidate because it is not only expressed in iridoblasts and iridophores, but also in many different types of cells, including premigratory NCCs, cells of the ICM, the aorta and, later in development, cells of the swim bladder epithelium. *ltk* is a promising candidate because its expression has not been detected throughout the premigratory NC domain, but only in a subset of NCCs, postulated to generate specified iridoblasts (Lopes et al. 2008). However, transient *ltk* activation has been hypothesised to occur in premigratory NCCs likely to correspond to the oligopotent progenitor, chromatoglioblast (Nikaido et al., in prep.), while the gene is also known to be expressed in the notochord (Lopes et al. 2008; this work). Finally, the previously indicated iridophore marker, *pnp4a*, was shown over the course of this project to also be expressed in other NC derivatives. Therefore, genes that are solely activated in the iridophore lineage have not been determined to date, although a recent study of the iridophore transcriptome revealed several novel candidates, which await further characterisation (Higdon, Mitra, and Johnson 2013).

Following determination of an iridophore-specific factor, its promoter can be used to drive Cre recombinase expression in a suitable transgenic line, which will allow for permanent fluorescent labelling of iridophore precursors and their derivatives. Thus, FACS of developing iridophores will allow for quantitatively determining gene expression levels in appropriate cell populations throughout development.

3.4.4.3. Parameter choice

For simulating model A, physiologically relevant parameter constants, namely dissociation constants, maximum gene expression rates and degradation rates, were derived from the published literature. However, in the case of dissociation constants, values had to be adjusted in order to be able to successfully simulate activation of gene expression. It remains to be determined whether altering the dissociation constant values is appropriate, or whether otherwise modifying chosen parameters, for instance by

decreasing the assigned degradation rates, would prove equally effective and, perhaps, more suitable. Furthermore, the choice of maximum gene expression rate might require adjustments. Specifically, in this study, this parameter value was based on experimental data that concern mRNA production. However, this rate may benefit from taking into account the rate of protein production, a process which is implied but not explicitly described in the mathematical model. Notably, recently published work shed light on absolute protein synthesis rates (G.-W. Li et al. 2014), which could prove useful by rendering parameter choice more physiologically relevant.

Overall, it is important to note that the currently adopted parameter constants were, to a certain extent, chosen arbitrarily although they reflected physiologically relevant values. Furthermore, literature-derived parameter values had to be modified to meet the needs of the model. Important next steps will involve further computational tests aiming to assess the sensitivity of the model to parameter alterations. Ultimately, absolute quantitation of the expression levels of genes in the cell population of interest at sequential developmental stages will allow for fitting of the model outputs to the data, thus estimating parameters and producing simulations with a higher degree of predictive power.

3.4.5. Summary

In this chapter, the preliminary GRN governing iridophore development was constructed using experimental approaches. This network aimed to describe the molecular basis of the transition from a NC progenitor towards a mature iridophore. The preliminary model A was mathematically modelled using a system of ODEs, which was solved numerically using MATLAB software. Computationally simulating model A generated predictions regarding how gene expression dynamics progress over time in the model. These predictions were subsequently compared to previously determined patterns of gene expression, in order to evaluate their consistency with known biology. In chapter 4, model A predictions are experimentally evaluated and testable hypotheses are drawn to eliminate observed inaccuracies. This leads to generating model B. The process of mathematical modelling and experimental testing is repeated to derive models C and D, ultimately aiming to generate a GRN, the mathematical model of which predicts established gene expression dynamics with a higher degree of accuracy.

4

Iterative improvements of the iridophore GRN

4.1. Introduction

Chapter 3 outlined the process of generating the preliminary GRN guiding iridophore specification from NCCs (model A), using experimental methods. It was demonstrated that *sox10* was essential for initiating *tfec* expression in specified iridoblasts, and that *tfec* and the RTK *ltk* maintained each other's expression via a positive feedback loop. Intriguingly, both *tfec* and *mitfa* were found to upregulate the iridoblast marker, *pnp4a*. This network structure closely resembled that of the previously established melanophore GRN (Greenhill et al. 2011). In that network, *sox10* was shown to upregulate the melanocyte master regulator, *mitfa*, which belongs to the same gene family as *tfec*, and *mitfa* was reported to maintain its own expression in specified and differentiating melanophores via a positive feedback loop. At the time, the factor with which *mitfa* formed the positive feedback loop remained unknown, but it is hypothesised that it corresponds to the RTK *kit* (A. Lapedriza, PhD thesis). *mitfa* was shown to activate melanophore differentiation genes, such as *dct*, *tyr* and *tyrp1*, in the same way that this study suggested *tfec* upregulates *pnp4a* expression. To describe the intergenic relationships in the iridophore network, a mathematical model was generated using a system of ODEs, which was used to predict how gene expression dynamics progressed over time in the developing iridophore lineage.

This chapter focuses on biologically testing the accuracy of the preliminary GRN and details the iterative process of mathematical modelling and experimentation (Greenhill et al. 2011), through which the preliminary iridophore GRN was developed. Guided by qualitative predictions from the mathematical models, testable hypotheses were drawn in order to improve the network. Such hypotheses were, first, theoretically evaluated by computationally implementing the mathematical model and then experimentally tested using classical genetics techniques, namely *in situ* hybridisation and mRNA overexpression studies, as well as the newly developed RNAscope technology.

Additionally, novel *tfec* mutant lines were generated using the recently widely adopted CRISPR/Cas9 technology for directed mutagenesis, in order to test corresponding loss of function predictions of the model.

Three iterations of modelling and experimentation led to proposing that (i) *pnp4a* activation depended on the presence of combinations of factors, including *sox10*, *tfec* and *mitfa*, (ii) a *sox10*/*tfec* positive feedback loop, likely mediated by *ltk*, was in place and (iii) an iridophore-specific factor R functioning as a repressor of *mitfa* was required. To rigorously assess the predictive power of models and reinforce conclusions, computational tools were developed, namely a Monte Carlo sampling algorithm.

4.2. Generation of *tfec* mutants using CRISPR/Cas9 mutagenesis

4.2.1. Generation *tfec* mutants

In order to evaluate the predictions of the model in the loss of *tfec* function context, it was important to generate *tfec* mutant lines, as the morpholino technology is often associated with non-specific effects (Eisen and Smith 2008). The CRISPR/Cas9 system was employed to induce site-directed mutagenesis in the open reading frame of *tfec* (Sander and Joung 2014). To generate *tfec* mutants, WT fish of the WIK genetic background were co-injected with CRISPR guide RNA (gRNA) and Cas9 mRNA. The gRNA sequence targeting *tfec* was designed by Dr J. Lister, whose group first investigated the effects of loss of *tfec* function (J. Lister, unpublished data; refer to section 2.2.2). At first, approximately 70 pg CRISPR gRNA and 370 pg Cas9 mRNA (1:5 ratio), were injected directly into the first (flat) cell during the first 10 min of development. At 2 dpf, 110 injected embryos were screened and no iridophore phenotype was observed. Death rates were not increased between uninjected embryos and injected siblings, and no significant abnormalities were noted (data not shown).

The experiment was repeated, this time injecting approximately 280 pg of CRISPR gRNA and 700 pg Cas9 mRNA per flat cell (1:2.5 ratio). The percentage of dead embryos did not increase between uninjected embryos and injected siblings ($p = 0.32$; Table 4.1). Furthermore, only 5 out of approximately 300 live embryos at 2 dpf presented with abnormalities. Together, these data suggested that the *tfec* CRISPR induced minimal non-specific effects, even when injected at relatively high dosages. On average, 17% of live embryos injected with *tfec* CRISPR/Cas9 showed complete loss of iridophores and a potential effect on melanophore morphology (Table 4.1; Fig. 4.1). For the phenotype to manifest, both *tfec* alleles in the majority, if not in all, of NCCs had been efficiently targeted, which led to the assumption that many somatic cells, likely also germline cells, each carried one or two *tfec* mutant alleles.

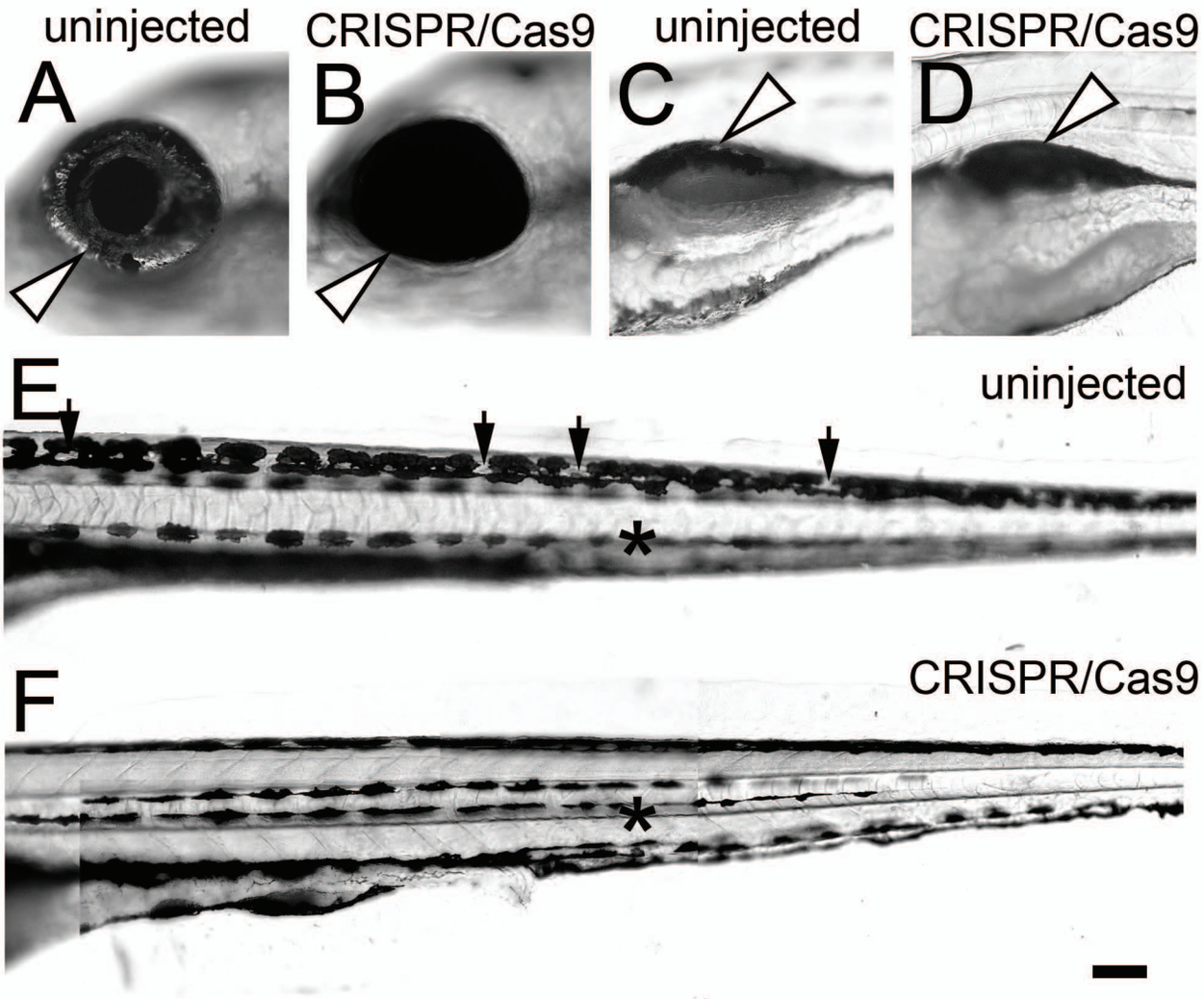


Figure 4.1. Iridophores are eliminated in embryos injected with tfec CRISPR/Cas9. Live imaging shows lack of iridophores over the eye (A, B; arrowheads), on the lateral patches (C, D; arrowheads) and along the dorsal stripe (E, F; arrows) of embryos injected with *tfec* guide RNA and Cas9 mRNA, compared to their WT siblings. Asterisks indicate melanocytes along the lateral stripe that looked morphologically different in injected embryos (E, F). Lateral views, oriented with the head to the left. Scale bars: 100 μm.

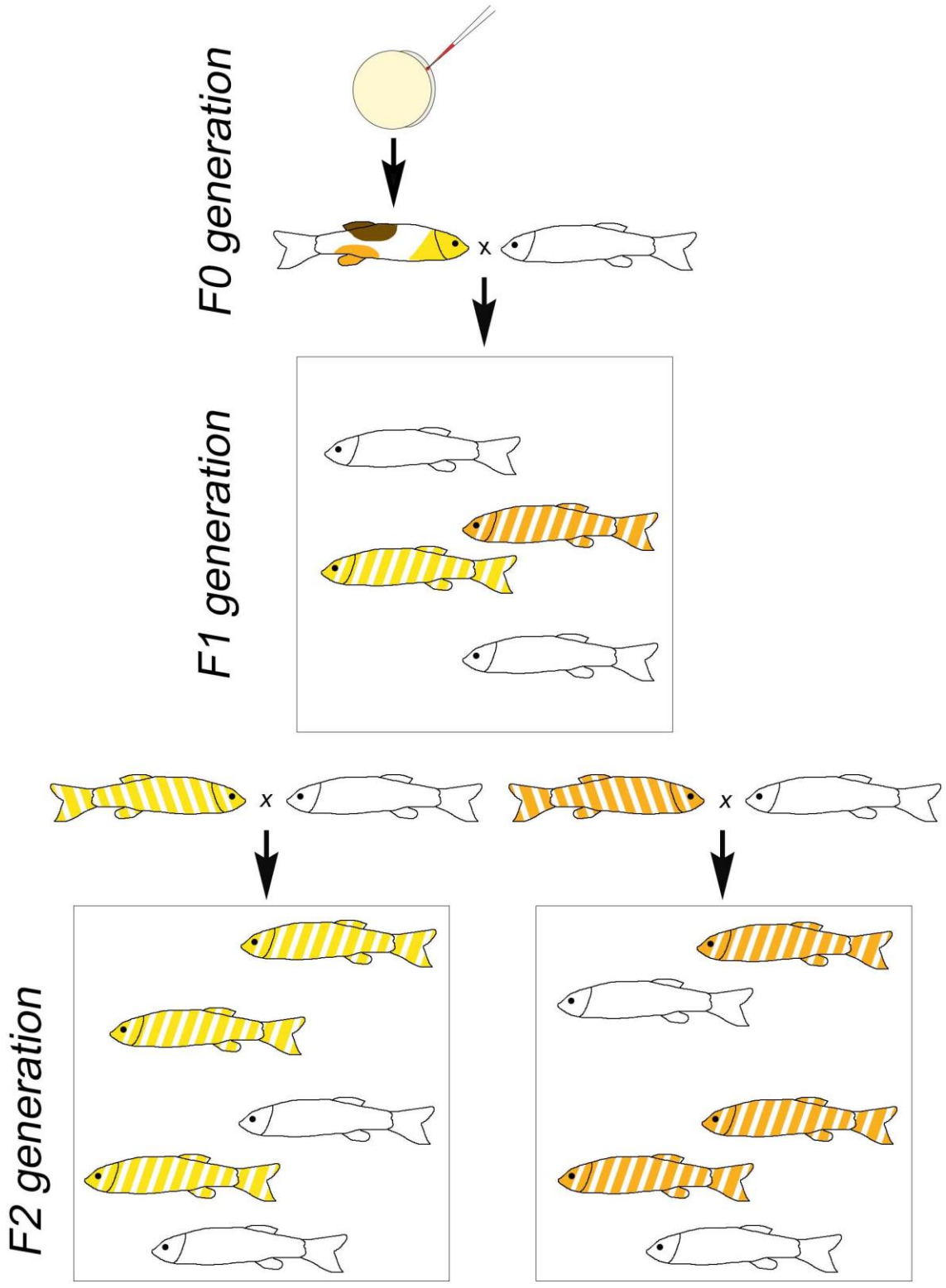

F0 generation
x
F1 generation
x
x
F2 generation

Illustration 4.1. Generation of mutant lines using the CRISPR/Cas9 system for site-directed mutagenesis. To generate a clean mutant line, WT embryos at the flat-cell stage (first 10 min of development) are injected with a mix of CRISPR guide RNA and Cas9 mRNA. The injected fish (F0) are screened for successful mutagenesis and raised to adulthood (approximately 3 months). Individuals of the F0 generation are likely chimeric (each colour represents arbitrarily chosen groups of cells with distinct mutations in the gene of interest). To generate non-chimeric fish, identified F0 adults that transmit mutations through their germlines are outcrossed to WT fish (not coloured). The progeny (F1) is raised to adulthood (approximately 3 months). Amongst the progeny are WT fish and fish which inherited a single mutant allele from the chimeric founder (heterozygotes; coloured with white stripes). F1 fish derived from a single outcross (contained in a rectangle) may thus comprise heterozygotes for different mutant alleles. At least two mutant carriers are identified, the corresponding mutations are determined by genomic DNA sequencing and they are outcrossed to WT fish. The progeny (F2) is raised to adulthood (approximately 3 months). F2 fish from each outcross comprise 50% carriers of a single mutation and 50% WT fish.

Table 4.1. Lethality and iridophore phenotype assessment in CRISPR-injected fish compared to WT siblings. Percentages were derived from a representative experiment. Approximately 100 uninjected and 60 injected embryos from each pair were scored.

Progeny of pair	Uninjected (death %)	Injected (death %)	Loss of iridophores (% of live embryos at 3 dpf)
1	7.1 %	12.7 %	10 %
2	13.7 %	19 %	11 %
3	21.4 %	27.9 %	27 %
4	9.8 %	6.2 %	14 %
5	16.4 %	22.8 %	2 %
6	75 %	69.5 %	38 %
Average	**23.9 %**	**26.3 %**	**17 %**

Injected embryos (F0) were raised to adulthood and screened for germline transmission of *tfec* mutant alleles. Several chimeric individual with partial loss of adult iridophores were observed amongst this generation, but none with complete loss of the lineage, suggesting that extensive loss of *tfec* function is likely to be deleterious. Pairs of injected F0 adults were incrossed and several individuals were identified that transmitted *tfec* mutant alleles through their germlines with high frequency. The phenotype of complete lack of iridophores was present in:

- 80-90% of embryos derived from 2 out of 6 incrossed pairs
- 50-60% of embryos from 2 out of 6 incrossed pairs
- 20% of embryos from 1 incross
- 0% of the progeny of 1 incross

Adults that transmitted *tfec* loss of function alleles with high frequency were outcrossed to WT fish (WIK background) and the progeny (F1) was raised to adulthood. Fish of the F1 generation were predicted to no longer be chimeric, meaning that each raised adult was either homozygous WT or heterozygous carrier of a single mutated allele. It was, however, likely that a single individual from the F0 transmitted different mutated alleles, and therefore tanks produced from a single F0 founder were likely to contain a mixture of heterozygotes for different *tfec* mutant alleles. To generate clean *tfec* mutant lines, F1 adults were incrossed and founders carrying *tfec* mutations were identified. These founders were again outcrossed to WT fish (WIK background) and the progeny (F2) was raised to adulthood. Different *tfec* mutant lines, derived from different F1 founders comprised individuals with the same mutation. Illustration 4.1 graphically describes the process of generating mutant lines using the CRISPR/Cas9 system.

4.2.2. Isolation of *tfec* mutant lines

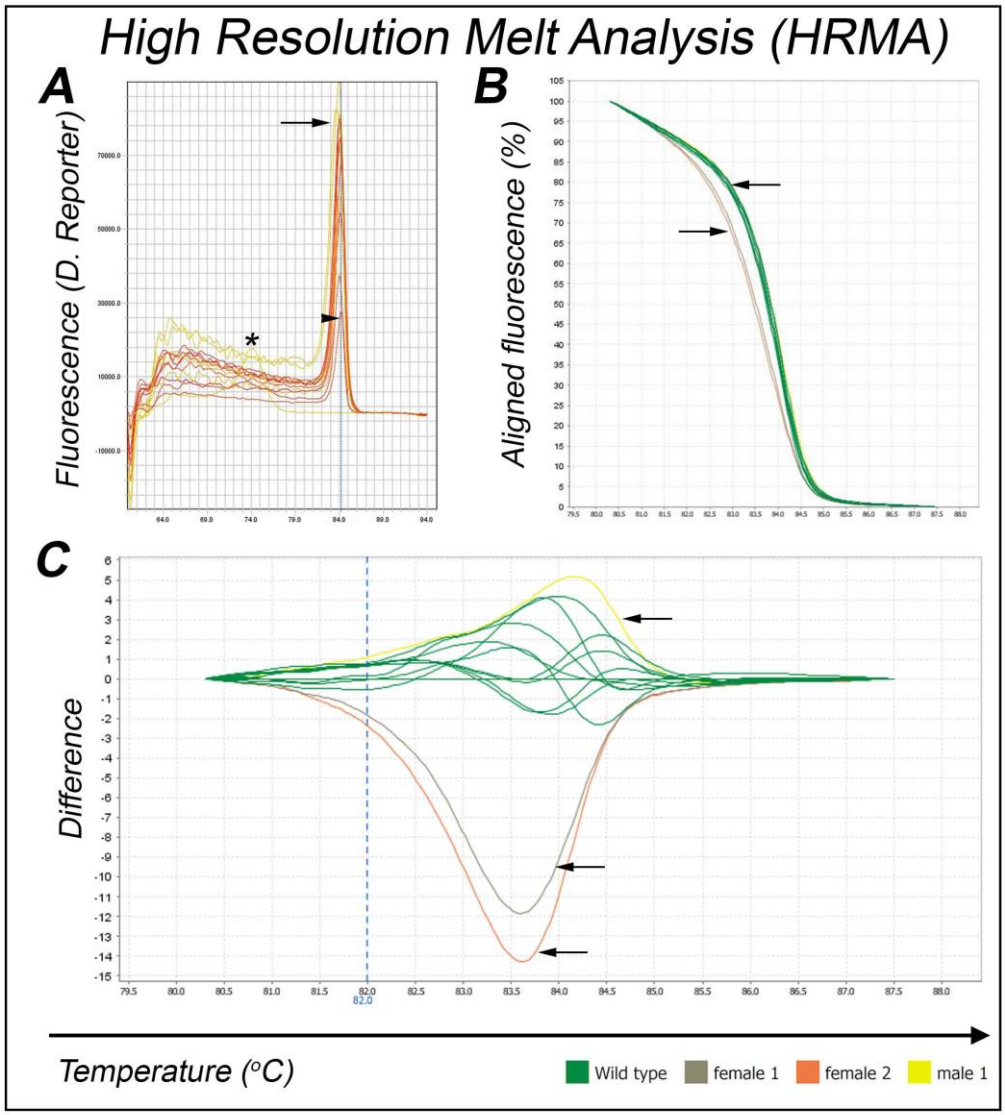
High Resolution Melt Analysis (HRMA) was employed in order to molecularly genotype *tfec* mutants. HRMA is a post-PCR analysis method designed to identify variations in DNA sequences, ranging from SNPs to extensive insertions and deletions ('in-dels') commonly generated by the CRISPR/Cas9 system. To perform HRMA, primers were designed to amplify the *tfec* genomic region flanking the CRISPR target sequence (refer to section 2.3.7 and to Appendix table II.2). Adult fish were swabbed to isolate genomic DNA (refer to section 2.3.5) and q-RT PCR was performed (refer to section 2.3.7). q-RT PCR results (Fig. 4.2 A, melt curve derivative) were analysed using HRMA software (Thermo Scientific). The software amplified very small differences between the melting temperatures of amplicons from different samples, which were visualised through the aligned melt curve plot (Fig. 4.2B). Finally, the algorithm generated a difference plot, which magnified the differences between reference control samples (WT or untreated) and unknown samples to be genotyped (Fig. 4.2 C).

To confirm the reliability of this genotyping methodology, HRMA was first performed using genomic DNA derived from 10 known WT adults (WIK background, 'reference' samples) and 3 F1 adults ('unknown' samples). The latter had been sequenced (Fig. 4.2 D) prior to raising the F2 generation, thus confirming two females to be mutation carriers, and one male to be homozygous WT. The two females were used as founders of two *tfec* mutant lines (F2 generation). HRMA correctly identified females 1 and 2 as *tfec* mutants (Fig. 4.2 B,C), and male 1 as WT (Fig. 4.2 C).

Once F2 adults were available, HRMA was employed to screen for carriers of *tfec* mutant alleles (data not shown). Identified individuals were incrossed to confirm that 25% of the progeny lacked iridophores, thus accounting for potential HRMA artefacts, which could result in false positives. For instance, low amplification efficiency of certain samples during q-RT PCR could generate downward deflecting difference curves even if the WT sequence was present in these samples. *tfec* mutation carriers from the female 2-derived stock (3 pairs) and from the female 1-derived stock (1 pair) were validated. All experiments described in the following sections were conducted using the former stock.

F2 identified carriers were incrossed and phenotypically WT embryos and mutant siblings were sequenced. The mutation causing loss of *tfec* function in the female 2 stock was shown to be a six nucleotide deletion:

WT: GGACGATCCTCAAGGCCTCGG
Mut: GGACGATCCTCAA----------GG


High Resolution Melt Analysis (HRMA)
A
Fluorescence (D. Reporter)
B
Aligned fluorescence (%)
C
Difference
Temperature (°C)
Wild type
female 1
female 2
male 1

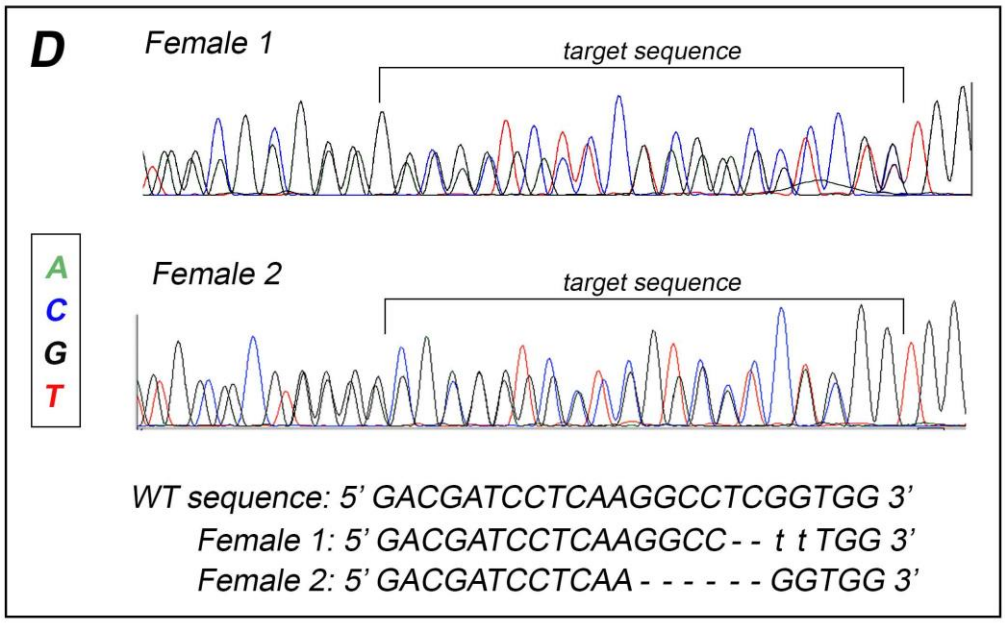

D
Female 1
target sequence
A
C
G
T
Female 2
target sequence
WT sequence: 5' GACGATCCTCAAGGCCTCGGTGG 3'
Female 1: 5' GACGATCCTCAAGGCC - - t t TGG 3'
Female 2: 5' GACGATCCTCAA - - - - - - GGTGG 3'

Fig. 4.2

Figure 4.2. Genotyping *tfec* mutants by High Resolution Melt Analysis (HRMA). (A) Melt curve (derivative reporter graph) generated by q-RT PCR. Clean, sharp peaks (arrow) are required for HRMA analysis. Peaks of different height (arrowhead) indicate lower amplification efficiency. In this assay, primer dimers (asterisk) appeared only in the negative control did not interfere with the peak of the melt curve. (B) Aligned melt curve plot. WT (green) and heterozygous mutant (grey) melt curves are clearly distinguishable. (C) Difference plot. Downward deflecting curves indicate mutagenesis and correspond to known mutation carriers (arrows, grey and orange lines). The known WT male (arrow, yellow line) clustered together with WT reference samples (green lines). (D) Sequencing results for female 1 and female 2. The CRISPR target sequence is indicated on the chromatograms. Sequencing was conducted using the reverse PCR primer (from the 3' end of the sequence). In female 1, four nucleotides were deleted (TCGG) and replaced by two nucleotides (tt). In female 2, six nucleotides were deleted (GGCCTC).

tfec encodes a bHLH TF. The identified in-frame mutation resulted in deletion of two consecutive amino acids (alanine and serine) from the C-terminal helix domain of Tfec (Fig. 4.3 C). This domain features the leucine zipper domain of the protein. Sequence alignments between the zebrafish Tfec sequence and Tfec and Mitf sequences from other species indicated that these residues are widely conserved and likely important for MiT TFs (Fig. 4.3 A,B). Whilst the alteration did not cause premature truncation and did not affect the DNA-binding domain of the TF, it likely affected either the structure of the Tfec monomer, or its ability to dimerise.

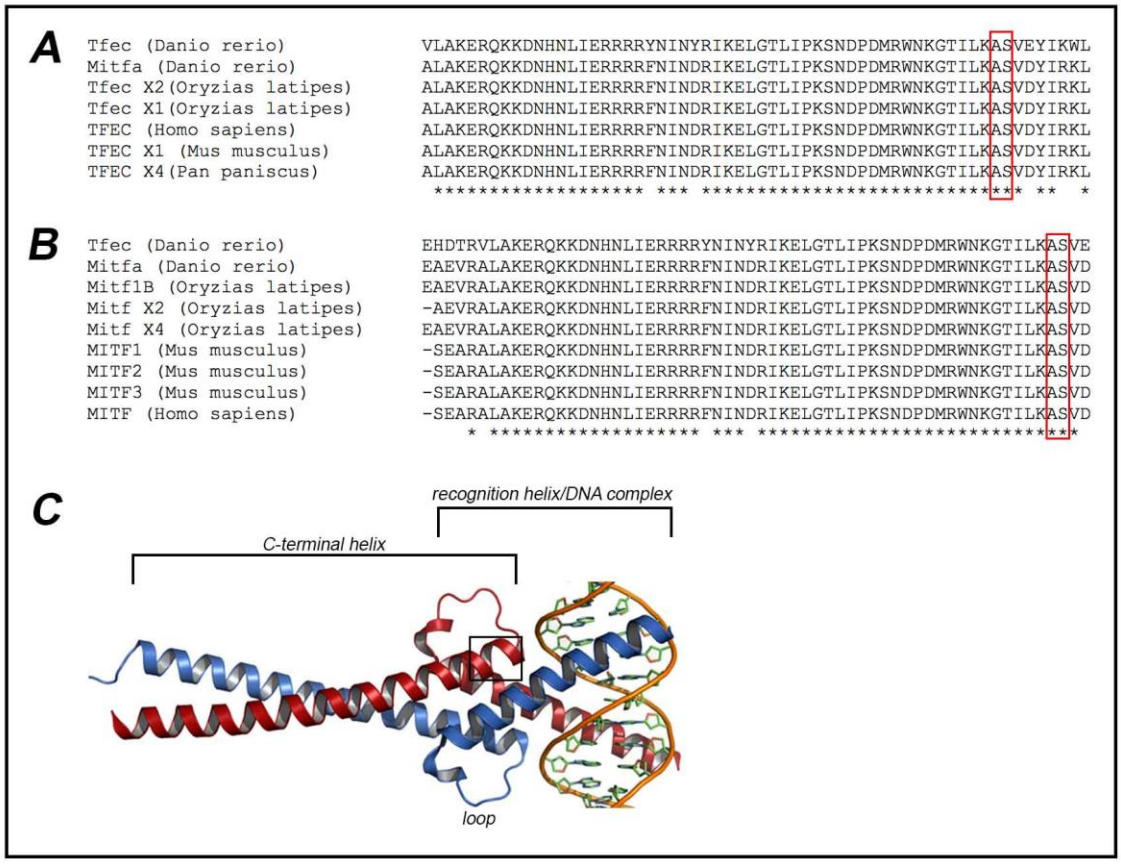


Figure 4.3. The deleted amino acid residues are conserved amongst MiT factors. Sequence alignment between (A) zebrafish, medaka, human, mouse and primate Tfec sequences and (B) zebrafish Tfec and medaka, mouse and human Mitf sequences shows conservation of the alanine and serine residues (AS) deleted in tfec mutants. (C) Structure schematic of a typical bHLH TF dimer binding to DNA. The boxed region at the start of the C-terminal helix is affected by the mutation. Figure adapted from online resources (http://slideplayer.com/slide/638089/).

4.2.3. Live phenotypes

tfec mutant carriers from the female 2 stock were incrossed and live embryos were observed. At 30 hpf, delay and potential reduction of melanogenesis was noticed in a subset of embryos derived from different crosses (Fig. 4.4 A,B). In the affected embryos, the eye was not melanised and very few melanophores occupied the dorsal trunk (Fig. 4.4 B). In comparison, phenotypically WT siblings demonstrated strongly pigmented eyes and noticeably more melanophores along the length of the dorsal trunk until the anterior tail region (Fig. 4.4 A). Because melanisation is first detected in the eyes and the anterior trunk, prior to the remaining trunk regions being occupied by pigment cells, these results suggested a general reduction in melanophore numbers in presumptive *tfec* mutants. Furthermore, in affected individuals, small numbers of melanised cells were observed along the medial and lateral pathways, and only towards the anterior of the embryo, while in presumed WT or heterozygous siblings melanophores occupied migratory pathways until the middle of the trunk (Fig. 4.4 A,B). This suggested that melanogenesis might be delayed in these embryos. Individuals that displayed the observed phenotypes were isolated from phenotypically WT siblings and monitored in the following days when differentiated iridophores were present.

All embryos that appeared phenotypically WT at 30 hpf displayed the normal pigmentation pattern at 4 dpf (Fig. 4.4 C,D). Iridophores were present along the dorsal, ventral and yolk sac stripes, on the LPs and overlaying the RPE. Melanophores were arranged in four stripes: the dorsal, lateral, ventral and yolk sac stripes. They also covered the eye, the dorsal region of the head, where a few iridophores were also located, and the dorsal yolk sac. In contrast, all presumptive mutant embryos displayed a complete lack of iridophores and a potential increase in melanophores, which has not yet been quantified (Fig. 4.4 E,F). Specifically, iridophores which normally occupied the dorsal, ventral and yolk sac stripes were absent, with very few or no detectable escapers. Moreover, the dorsal head, LPs and the layer overlaying the RPE lacked iridophores. A mild increase in melanophores occupying the head and yolk sac of the embryos was observed, although further quantitation is required to confirm this phenotype.

To confirm that the observed phenotypes correlated with the described mutation in the *tfec* locus, genomic DNA from five isolated presumptive mutant embryos and from five WT siblings was extracted. The region of interest was amplified by PCR (refer to Appendix table II.4) and sequenced. All five embryos were found to be homozygous mutants, while the sequenced WT siblings were heterozygous carriers of the mutant allele (data not shown). Finally, at 5 dpf a severe defect in swim bladder formation was noticed. Embryos lacking iridophores failed to inflate their swim bladders, indicating a role of *tfec* in development of the organ (data not shown).

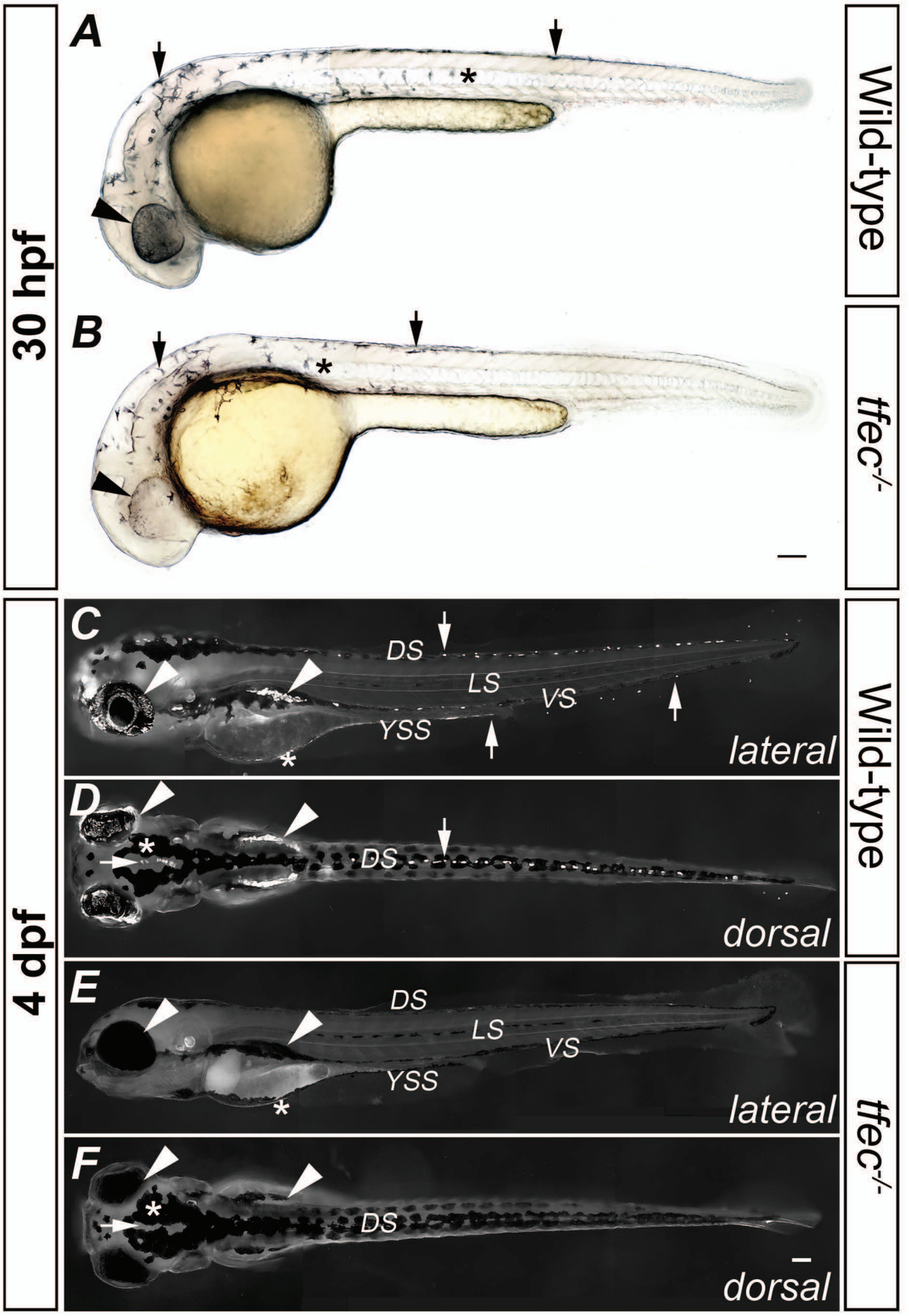

30 hpf
4 dpf
Wild-type
tfec-/-
Wild-type
tfec-/-
A
B
C
D
E
F
DS
LS
VS
YSS
lateral
dorsal
lateral
dorsal

Fig. 4.4

Figure 4.4. Development of both melanophores and iridophores is affected in *tfec* mutants. Live imaging of genotyped WT and mutant embryos at 30 hpf (A, B) and at 4 dpf (C-F). At 30 hpf, melanophores along the dorsal trunk (arrows), migratory pathways (asterisks) and covering the RPE (arrowheads) are reduced in *tfec* mutant embryos (B), compared to their WT siblings (A). At 2 dpf, lateral (C) and dorsal (D) views of a WT fish and of a *tfec* mutant (E, F, respectively) show loss of iridophores along the dorsal, ventral and yolk sac stripes (vertical arrows), the dorsal surface of the head (horizontal arrows), the lateral patches and overlaying the RPE (arrowheads). Melanocytes (asterisks) occupy the ventral yolk sac and appear increased on the head of the mutant embryo (E, F), but not of the WT (C, D). Oriented with the head to the left. Scale bars: 100 μm.

4.3. Experimental evaluation of model A predictions

Chapter 3 outlined the generation of model A and the process of mathematically modelling intergenic interactions within that model. The model provided qualitative predictions of how the expression dynamics would progress during development of the iridophore lineage in the WT and in loss of gene function contexts (Fig. 3.16). In this section, experimental tests to determine the degree of consistency of model A predictions with *in vivo* biology are presented. For these tests, existing WT and mutant zebrafish lines, as well as the newly established *tfec* mutant lines described in the previous section were instrumental.

4.3.1. Evaluating the WT context

Based on data from published literature, *sox10* expression was manually set to be activated between 18 hpf and 22 hpf, and to decline following 22 hpf (refer to section 3.3.2). Nevertheless, the effects of alternative *sox10* expression dynamics were tested theoretically by manually altering the Sox10 degradation rate in the equation describing the gene's behaviour following 22 hpf (Fig. 4.5). This resulted in Sox10 being maintained for longer, even until 100 hpf, when degradation rates were small, or being rapidly eliminated when degradation rates were large. It was demonstrated that, regardless of the behaviour of Sox10, the concentrations of Ltk, Tfec and Pnp4a qualitatively remained the same, reaching positive plateau levels at differentiation stages. However, it was observed that prolonged maintenance of Sox10 resulted in maintenance of *mitfa* expression in the iridophore lineage. It was assumed that the melanophore gene *mitfa* is not expressed in differentiated iridophores, hence this predicted behaviour was regarded as biologically inaccurate. Based on these tests and data regarding *sox10* dynamics in the NC-derived melanophore lineage (Greenhill et al. 2011), the Sox10 degradation rate following 22 hpf was set to $d_S = 0.175\,1/h$. Using this rate, Sox10 dropped to undetectable levels by approximately 48 hpf (Fig. 4.5), in line with its behaviour in the melanophore lineage (Greenhill et al. 2011).

Having established a seemingly appropriate behaviour of the Sox10 input, the biological accuracy of generated predictions in the WT context, when all genes are able to activate their downstream targets, was assessed (Fig. 3.16; Wild type simulation).

In this study, as well as in previously published work, expression of both *tfec* (Fig. 3.1; Lister et al. 2011) and *ltk* (Fig. 3.6; Lopes et al. 2008) was found in positions of mature iridophores. Moreover, both *pnp4a* transcript (Fig. 3.11; Curran et al. 2010) and Pnp4a protein (Curran et al. 2010) have been shown to co-localise with mature iridophores. Therefore, the predicted maintenance of these marker genes throughout iridophore development was considered biologically accurate.

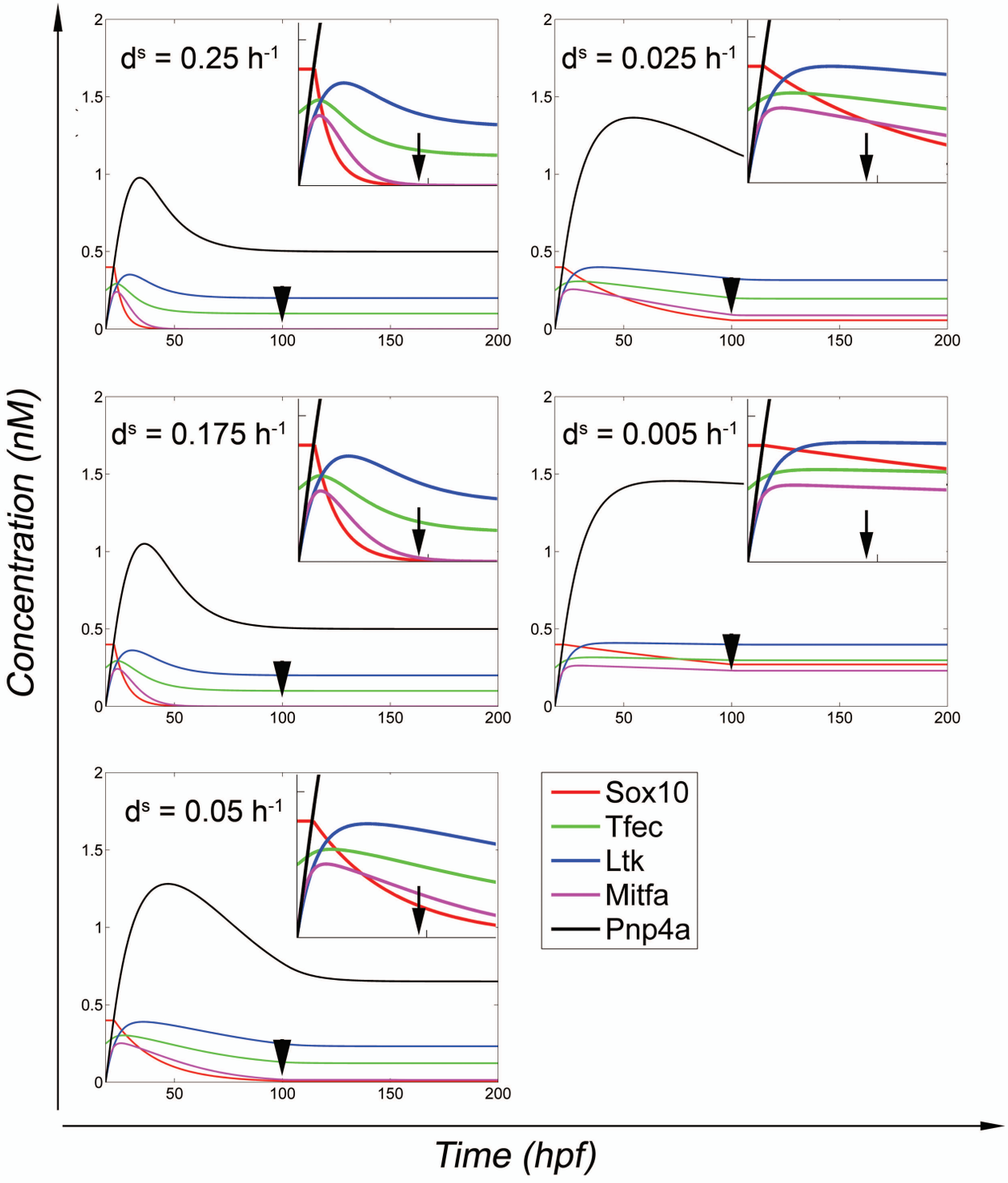


Figure 4.5. Theoretical testing of alternative *sox10* expression dynamics in model A. Different Sox10 degradation rates (d^S) were implemented. Mitfa concentration (arrowheads) is closely correlated with Sox10 concentration. At t=48 hpf (arrows) *sox10* expression should no longer be detectable (Greenhill et al. 2011). According to this criterion d^S=0.175 h^{-1} was the most appropriate degradation rate for Sox10 after 22 hpf.

The model's prediction regarding behaviour of the melanophore gene, *mitfa,* in the iridophore lineage was next assessed. The gene has been shown to be expressed in the melanophore lineage and to be essential for its development (J A Lister et al. 1999), but expression in iridophores has not been reported to date. Therefore, the predicted downregulation of *mitfa* as iridophore development progresses (Fig. 3.16; Wild type simulation) was deemed biologically accurate. In order to validate the predicted initial upregulation of *mitfa* in early iridophore progenitors, which is a direct consequence of *sox10* presence at this stage, co-expression of *mitfa* with the iridoblast marker, *tfec*, was investigated between 18 hpf and 48 hpf using the RNAscope technology (refer to section 2.6). RNAscope allowed for fluorescent visualisation of *mitfa* and *tfec* transcripts simultaneously in embryos fixed at each of the stages of interest. Therefore, *tfec* positive cells, likely corresponding to iridoblasts, could be scored for *mitfa* expression at each stage (Fig. 4.6; refer to section 2.6.4).

In accordance with results derived from *in situ* hybridisation experiments (Fig. 3.15 G,H), *mitfa* expression at 18 hpf emerged in the NCC domain of the anterior dorsal trunk (Fig. 4.6 A). At this stage, *tfec* expression was widespread across the dorsal NCC domain (Fig. 4.6 B) and the majority, if not all, of nuclei surrounded by *mitfa* fluorescent signal were located within this *tfec* expressing domain (Fig. 4.6 D). At 24 hpf, dorsally positioned and migrating melanoblasts prominently expressed *mitfa* (Fig. 4.6 E). As was indicated by the small number of the characteristic spots surrounding nuclei of dorsally located and migrating NC derivatives, *tfec* was being downregulated in most of these cells (Fig. 4.6 Fi). Towards the anterior trunk of the embryo, a small number of cells retained high levels of *tfec* expression, indicated by nearly solid staining of the cytoplasm (Fig. 4.6 F ii,iii). These cells presumably corresponded to specified iridoblasts and co-expressed *mitfa* (Fig. 4.6 H ii,iii). By 30 hpf, the majority of *tfec* positive cells, corresponding to dorsally located and migrating iridoblasts, were found to be *mitfa* negative, with very few cells of the less developed posterior trunk still expressing both transcripts (Fig. 4.6 I-L). At 36 hpf and 48 hpf co-expression of the two genes was not detected (data not shown), consistent with the view that *mitfa* is a melanophore marker and *tfec* is an iridophore marker.

Conclusively, experimental data supported the model's prediction that *mitfa* is upregulated in NCC derivatives undergoing initial stages of specification towards the iridophore lineage, identified based on prominent expression of *tfec*. Taken together, the expression dynamics of all genes were accurately predicted by model A in the WT context.

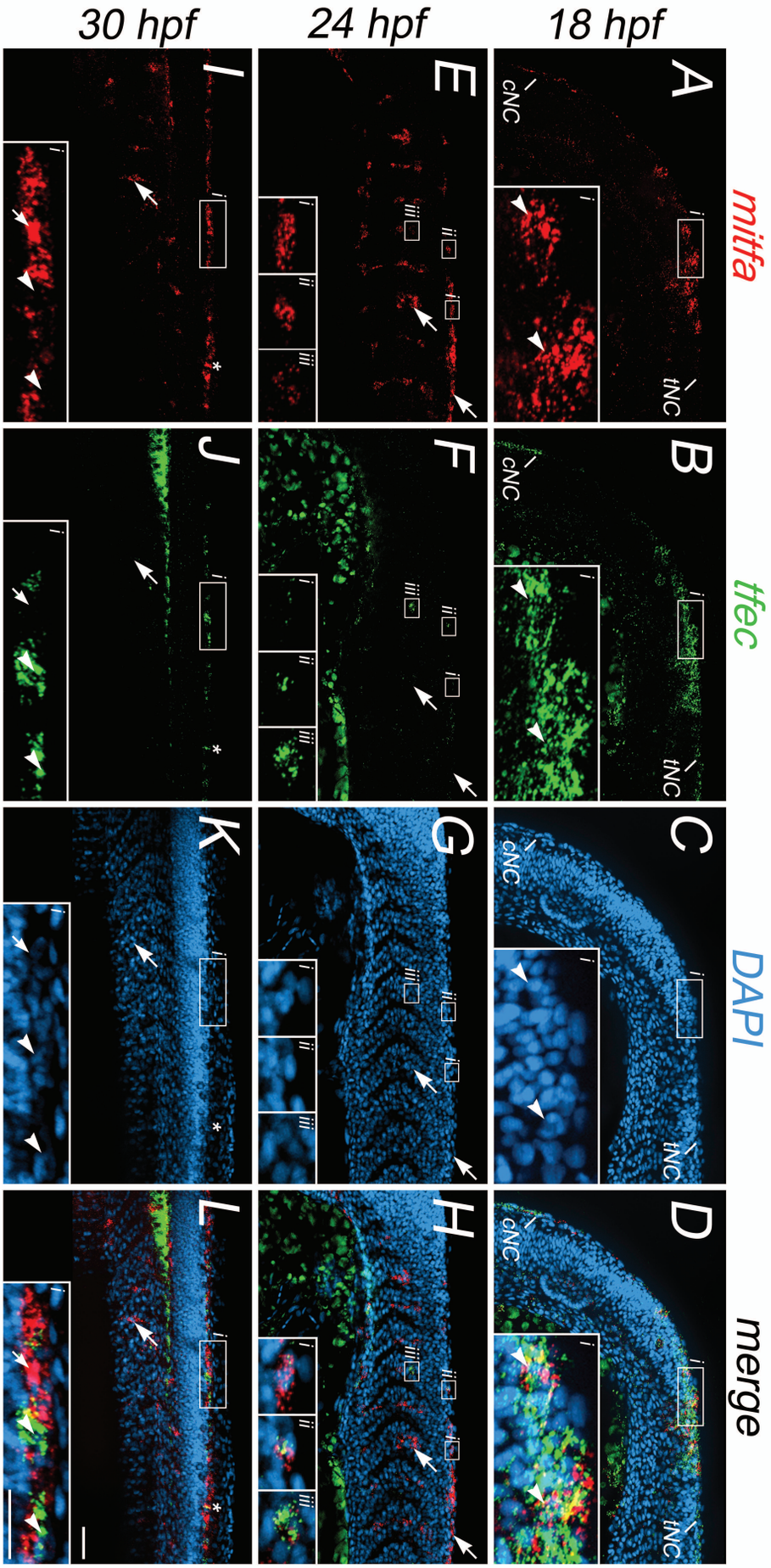

30 hpf
24 hpf
18 hpf
mitfa
tfec
DAPI
merge
A
B
C
D
E
F
G
H
I
J
K
L
cNC
tNC
i
ii
iii
iiii
*

Fig. 4.6

Figure 4.6. *mitfa* is expressed in early iridoblasts. Co-labelling of *mitfa* (red) and *tfec* (green) transcripts in WT embryos at 18 hpf, 24 hpf and 30 hpf using RNAscope. White squares indicate position of insets. At 18 hpf (A-D), *mitfa* expression emerges within *tfec* positive cNC and tNC domains. Prominent *mitfa* expression in anterior tNCCs (arrowheads) overlaps with *tfec* expression (i, arrowheads). At 24 hpf (E-H), *mitfa* labels dorsally located and migrating melanoblasts (arrows). *tfec* is expressed in low levels in several *mitfa* positive cells (i), and strongly maintained in very few NC derivatives, which likely correspond to specified iridoblasts and express *mitfa* (ii, iii). At 30 hpf (I-L), dorsally located (i, upward arrows) and migrating (downward arrows) melanoblasts express *mitfa* and dorsally located iridoblasts (i, arrowheads) express *tfec*. Very few cells of the posterior trunk show co-expression (asterisk). cNC, cranial neural crest; tNC, trunk neural crest. Lateral views, oriented with the head to the left. Scale bar: 50 μm. Insets: 25 μm.

4.3.2. Evaluating the loss of *sox10* function context

The next step in assessing the biological relevance of model A was to experimentally evaluate predictions in the loss of *sox10* function context, when the equations are such that Sox10 cannot activate its downstream targets (Fig. 3.16; *sox10*$^{-/-}$ simulation). In most aspects, the predicted gene expression changes fitted well with the known biology.

Firstly, according to model A, loss of *sox10* function did not result in any changes in *sox10* expression. This was considered biologically accurate, because *sox10* expression has been previously shown to persist in NCCs of *sox10* mutant embryos (Dutton et al. 2001). Secondly, it was predicted that the iridogenic genes *tfec* and *ltk* remained upregulated throughout development. These predictions were interpreted as biologically accurate, based on *in situ* hybridisation data suggesting that both genes were expressed in NCCs of the dorsal trunk in *sox10* mutants (refer to sections 3.2.2 and 3.2.3). Thirdly, *mitfa* was predicted to never become upregulated in *sox10* mutants, in line with previously published data (Dutton et al. 2001).

However, in one important aspect the model performed poorly. Specifically, *pnp4a* expression was predicted to persist, due to prolonged expression of *tfec* in the NCC domain. This prediction was experimentally tested by performing *in situ* hybridisation to detect *pnp4a* expression in samples which contained WT/heterozygous and homozygous *sox10*^{t3} mutant siblings. Surprisingly, in *sox10* mutants, *pnp4a* expression failed in the locations where *tfec* was prominently expressed (refer to section 3.2.2). Instead, in 13/45 24 hpf embryos examined over 2 independent experiments (most likely mutants according to Mendelian ratios; $0.5 < \text{p-value} < 0.7$), *pnp4a* transcript was eliminated from cNCCs and from ventrally migrating cells and was only expressed in relatively few trapped progenitors of the dorsal trunk NCC domain. The RPE domain remained unaffected in these embryos (Fig. 4.7 A,B).

By 30 hpf, in 8/39 assessed embryos (most likely mutants according to Mendelian ratios; 2 experimental repeats; $0.4 < \text{p-value} < 0.5$), dorsal trunk expression of *pnp4a* was restricted to very few cells and was detected at very low levels compared to presumptive WT siblings (Fig. 4.7 C,D insets). Moreover, migrating cells and LP progenitors were no longer stained in presumptive mutants. No differences were observed in the RPE region at this stage (Fig. 4.7 C,D). At 36 hpf, 12/40 embryos (most likely mutants according to Mendelian ratios; 2 experimental repeats; $0.3 < \text{p-value} < 0.5$) either bore a single escaper cell or completely lacked cells along the dorsal trunk. Staining in the eyes and in the LPs, consistently present in siblings, was eliminated and migrating cells were absent (Fig. 4.7 E-F). Finally, at 48 hpf 11/33 embryos assessed in 2 experimental repeats (most likely mutants according to Mendelian ratios; $0.2 < \text{p-value} < 0.3$) displayed on average 3 escaper cells along the dorsal stripe and no escapers along the ventral or

the yolk sac stripes. Expression in the LPs and overlaying the RPE was also severely affected in these embryos (Fig. 4.7 G,H).

Overall, experimental evidence suggested that model A inaccurately predicted *pnp4a* expression dynamics in the *sox10* mutant context. It became apparent that expression of *tfec* was not sufficient to upregulate *pnp4a* in the absence of *sox10*.

4.3.3. Evaluating the loss of *ltk* function context

In the loss of *ltk* function context, when Ltk was rendered incapable to initiate intracellular signalling that would result in upregulation of *tfec*, model A predictions reflected biological observations (Fig. 3.16; $ltk^{-/-}$ simulation). Previous analyses had demonstrated that *ltk* positive cells normally present in the premigratory NCC domain were reduced from 26 hpf and completely absent by 30 hpf in ltk^{ty82} mutant embryos (Lopes et al. 2008). Furthermore, as stated in section 3.2.5, *tfec* expression appeared unaffected in ltk^{ty82} mutant embryos between 18 hpf and 24 hpf. However, at 30 hpf, a significant reduction in the numbers of *tfec* positive cells was observed (Fig. 3.9). Together, these results suggested that the predicted upregulation and subsequent decline to undetectable levels of both Ltk and Tfec prior to iridophore differentiation were biologically accurate.

Similarly, Pnp4a concentration was predicted to initially rise and then decline later in development, following the trends of Tfec and Ltk. To experimentally evaluate this prediction, *pnp4a* expression was examined using *in situ* hybridisation in samples comprising WT embryos, ltk^{ty82} heterozygotes and homozygous mutants. At 24 hpf, *pnp4a* expression appeared unaffected in 78/81 assessed embryos (p-value < 0.001; 4 experimental repeats; data not shown). However, by 30 hpf, strongly labelled cells normally present along the dorsal trunk were reduced in 5/27 embryos (most likely mutants according to Mendelian ratios; 0.3 < p-value < 0.5; Fig. 4.8 A,B). Correlation of the presumed mutant phenotype with homozygosis of the mutated *ltk* allele was confirmed by genotyping all the embryos in the sample (refer to section 2.3.2; data not shown). This phenotype was observed in 2 independent replicates, although molecular genotyping was only carried out for one replicate. It should be noted that cells in the premigratory NCC domain of the dorsal tail were still detectable in mutant embryos, as well as migrating cells along the trunk and weakly expressing cells along the dorsal trunk (Fig. 4.8 A,B). These cells were also present in genotyped WT and heterozygote siblings and, based on *pnp4a* expression results in $mitfa^{w2}$ mutants at this stage (refer to section 3.2.6), likely corresponded to developing melanoblasts.

At 36 hpf, lack of cells strongly expressing *pnp4a* along the dorsal trunk and the migratory pathways was obvious in 7/28 examined embryos (most likely mutants according to Mendelian ratios; one experimental repeat; p-value > 0.95), owing to reduction in *mitfa*-

dependent *pnp4a* expression at this stage. These presumptive mutants were further distinguishable based on the noticeable reduction in staining associated with the RPE (Fig. 4.8 C,D). All embryos in the sample were molecularly genotyped to confirm that the observed phenotype manifested only in *ltk*^{ty82} homozygous mutants and was thus not due to experimental artefacts (data not shown).

Finally, at 48 hpf, the number of *pnp4a*-expressing cells normally present in iridophore locations was severely reduced in 7/28 assessed embryos (most likely mutants according to Mendelian ratios; one experimental repeat; p-value > 0.95; Fig. 4.8 E-H). In WT embryos, approximately 25 *pnp4a* expressing cells were scored along the dorsal stripe and between 25-28 cells along the ventral stripe. In embryos displaying the observed phenotype, 1 cell on average was detectable along the dorsal stripe and 2 cells on average along the ventral stripe. To verify that the observed phenotype was not due to experimental artefacts, representative embryos for both observed phenotypes were molecularly genotyped and their expected identities were confirmed (data not shown). The conclusion that lack of *pnp4a* positive cells results from loss of *ltk* function was reinforced by carrying out *in situ* hybridisation on samples containing *ltk*^{ty82} homozygous mutants at 60 hpf. Over 3 independent experimental repeats, 19/60 examined embryos presented with the described lack of *pnp4a* positive cells (0.2 < p-value < 0.3; data not shown).

It is worth noting that escaper cells in *ltk*^{ty82} mutants, as shown in figure 4.8 (F,H), were very strongly stained, similarly to cells in the WT, while escaper cells in *sox10*^{t3} mutants (Fig. 4.7 H,J) appeared much weaker than those in WT siblings, suggesting different roles for *sox10* and *ltk* in differentiation of the iridophore lineage. These results supported the prediction of model A that *pnp4a* first becomes activated and then declines to undetectable levels in the *ltk* mutant context.

mitfa expression dynamics were not experimentally assessed in *ltk* mutants. However, it has been reported that neither the number of melanophores nor the number of *dct* positive melanoblasts were significantly altered in *ltk*^{ty82} mutants compared to WT siblings (Lopes et al. 2008). Thus, *mitfa* expression was unlikely to be affected in this context, which was reflected by model A predictions.

Overall, predictions regarding gene expression dynamics in model A upon loss of *ltk* function accurately reproduced the experimentally observed dynamics in the iridophore lineage.

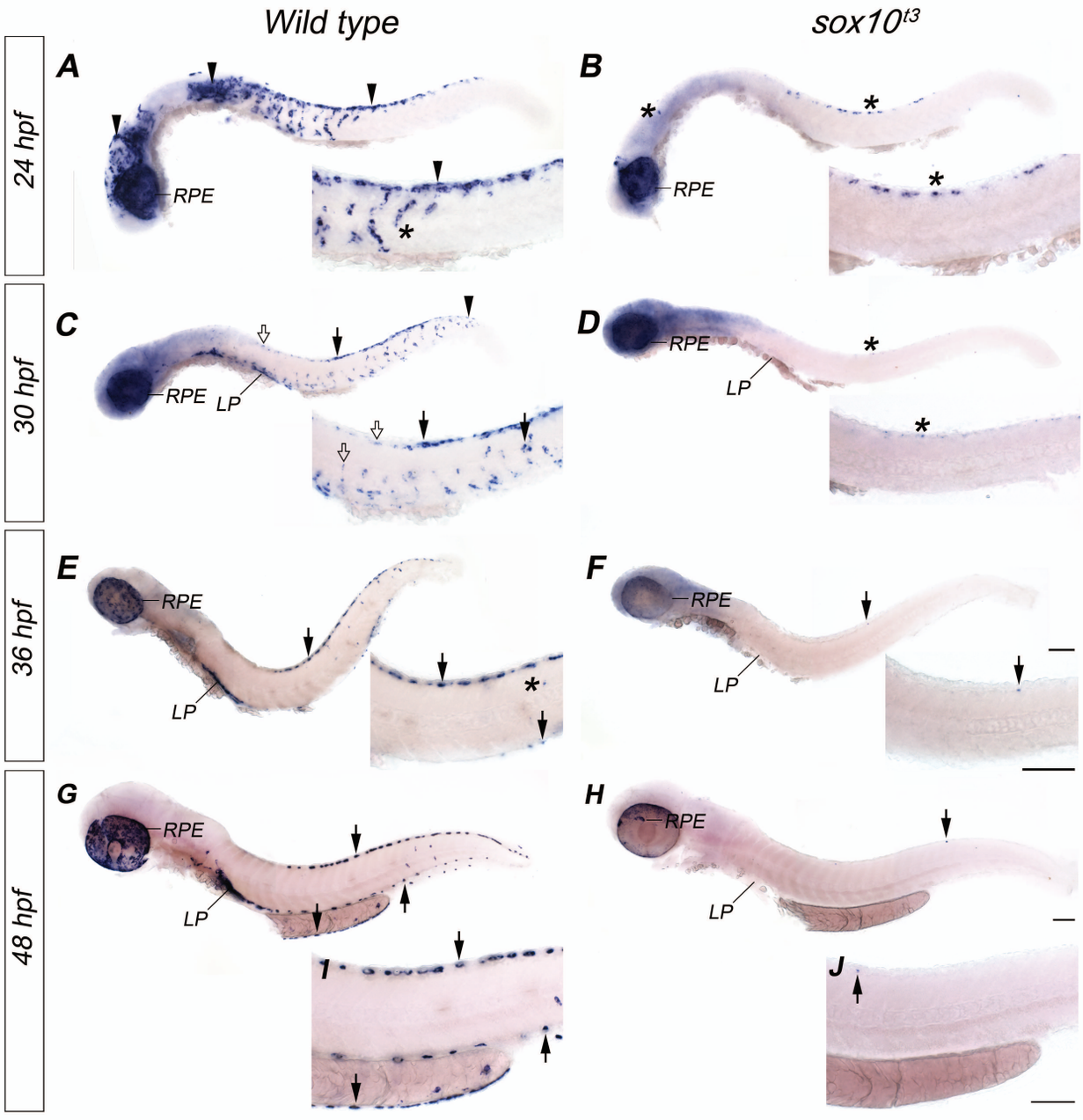


Figure 4.7. *pnp4a* expression fails in *sox10* mutants. *In situ* hybridisation results showing *pnp4a* expression in phenotypically WT (A, C, E, G, I) and presumptive *sox10^{t3}* mutants (B, D, F, H, J) at 24 hpf (A, B), 30 hpf (C, D), 36 hpf (E, F) and 48 hpf (G-J). In the WT, pnp4a is expressed in premigratory NCCs (arrowheads), migrating cells (A, E, asterisks), dorsal trunk iridoblasts (C, solid arrows), iridophores of the dorsal, ventral and yolk sac stripes (G, I, arrows) and weakly in putative melanophore lineage cells (empty arrows). Cells in the vicinity of the RPE and the LPs also express *pnp4a*. In mutants, *pnp4a* is expressed in very few premigratory NCCs (B, D, asterisks) and weakly in escaper iridophores along the dorsal trunk (F, H, arrows). RPE, retinal pigment epithelium; LP, lateral patches. Lateral views, oriented with the head to the left. Scale bars: 100 µm.

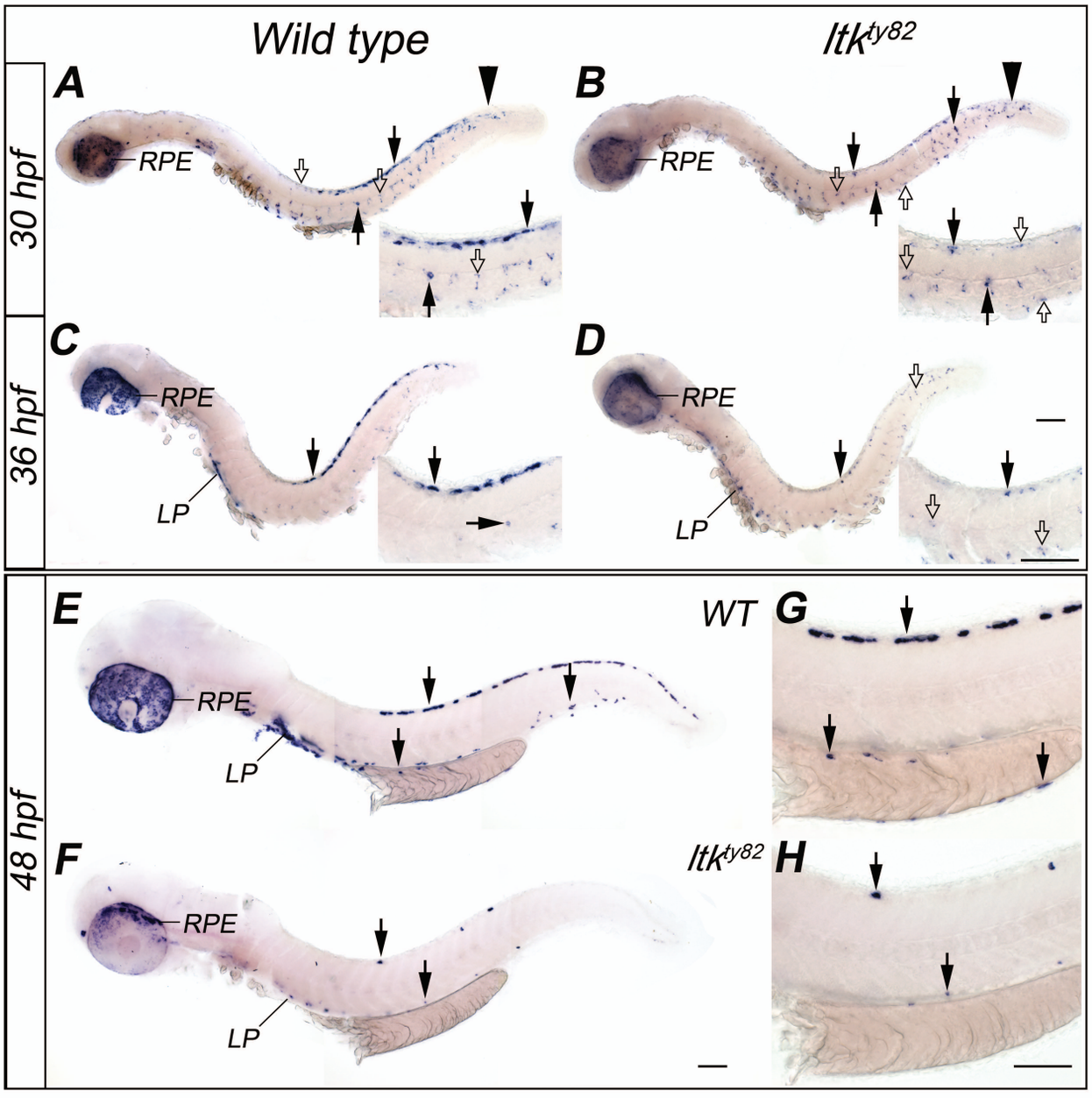


Figure 4.8. *pnp4a* expression declines during iridoblast specification in *ltk^{ty82}* mutants. *In situ* hybridisation results showing *pnp4a* expression patterns in genotyped WT and *ltk^{ty82}* mutant embryos at 30 hpf (A, B), 36 hpf (C,D) and 48 hpf (E-H). The premigratory NC domain (arrowheads) is unaffected in mutant embryos, while numbers of specified iridoblasts along the dorsal trunk (A-D; solid arrows) and of mature iridophores along the dorsal, ventral and yolk sac stripes (E-H; solid arrows) are decreased. Expression in the melanophore lineage (empty arrows) appears unaffected in ltk mutants (A-D). Expression in the vicinity of the RPE and in the developing LPs is severely reduced in mutants. RPE, retinal pigment epithelium; LP, lateral patches. Lateral views, oriented with the head to the left. Scale bars: 100 µm.

4.3.4. Evaluating the loss of *tfec* function context

The newly generated *tfec* mutant line (refer to section 4.2) was used to experimentally evaluate the model's predictions in the *tfec* loss of function context. It should be noted that results in this section were derived from a single experimental repeat, and were thus considered preliminary. All experiments performed at 24 hpf and 30 hpf were carried out using samples containing a mix of *tfec* homozygous mutants, heterozygous carriers and WT siblings obtained from incrossing identified mutation carriers (female 2 stock, F2 generation). Results at 2 dpf were first performed using embryos 25% of which lacked iridophores, obtained from a single F1 incross (Fig. 4.11; Fig. 4.12). It was thus likely that these embryos were carriers of two different *tfec* mutant alleles that resulted in complete lack of iridophores when homozygosed. These results were evaluated using 4 embryos per gene (3 which displayed the known WT iridophore phenotype, 1 with lack of iridophores) obtained from in-crosses of the pure *tfec* mutant stock (female 2 line, F2 generation). These data are not shown but replicated the results shown in figure 4.11.

In most aspects, model A predictions when *tfec* was rendered unable to trigger gene expression fitted well with *in vivo* biological evidence. Model A predicted that *ltk* never becomes upregulated upon loss of *tfec* function, when *tfec* was rendered unable to upregulate its target genes (Fig. 3.16; $tfec^{-/-}$ simulation). The *ltk* expression phenotype observed upon morpholino-mediated loss of *tfec* function indicated that, indeed, *ltk* was not expressed in the normal positions at 30 hpf (refer to section 3.25). *tfec* mutants were used to replicate this result at more time points. At 24 hpf, 6/26 embryos (most likely mutants according to Mendelian ratios; one experimental repeat; $0.8 < \text{p-value} < 0.9$) from a mix of *tfec* homozygous mutants, heterozygous and WT siblings showed reduction in the number of NCC domain progenitors located in the dorsal mid section of the trunk (data not shown). In presumptive WT siblings, there were approximately 20 cells in this position, arranged in a bilateral manner. *ltk* expression was also emerging in the dorsal eye region. In presumptive mutants, 6 cells on average were scored along the dorsal trunk, and the eyes were not stained (data not shown). The experiment requires repetition to confirm that the effect was not an experimental artefact, for example due to small variations in the developmental stage of individual embryos.

By 30 hpf, the effect of loss of *tfec* function on *ltk* expression was clearly discernible. 6/19 examined embryos (most likely mutants according to Mendelian ratios; one experimental repeat; $0.5 < \text{p-value} < 0.7$) lacked the majority of *ltk* positive iridoblasts along the dorsal trunk, the migrating pathways, the LP domain and the eye region (Fig. 4.9). The few cells still detectable in the affected embryos (0-2 along the dorsal trunk and approximately 5 within the migrating pathways) were faintly stained, indicating low levels

of *ltk* expression compared to cells in WT siblings. This observation is in line with the working hypothesis that *tfec* plays a role in sustaining *ltk* expression in the iridoblast.

At 2 dpf, 7/15 embryos (low probability of corresponding to mutants according to Mendelian ratios; one experimental repeat; 0.05 < p-value < 0.1) displayed severe reduction in *ltk* expressing iridophores along the dorsal, ventral and yolk sac stripes. The low p-value associated with this result is most likely a consequence of the small sample number (N = 15) potentially obscuring Mendelian ratios. Obtaining similar p-values even after the sample number increases would either indicate a non-Mendelian mode of inheritance of the mutant allele, or perhaps that heterozygous *tfec* mutants also present with defects in *ltk* expression. In the affected embryos, cells overlaying the RPE were also barely noticeable while both the LPs and the swim bladder epithelium were no longer expressing *ltk*. The few remaining escapers were stained very weakly compared to cells in WT siblings reinforcing the working hypothesis that *tfec* plays a crucial role in maintaining *ltk* expression levels (Fig. 4.11 A-D; Fig. 4.12).

Although *ltk* positive cells were still detectable in presumptive *tfec* mutants at 24 hpf, their reduced numbers might indicate that iridoblast-specific upregulation fails from early stages, as is suggested by the model's predictions. Further experiments are required to confirm whether an initial upregulation of *ltk* expression is followed by a rapid decline upon loss of *tfec* function, or whether *ltk* is never activated in iridoblasts, in line with model A predictions.

Model A predicted that, upon loss of *tfec* function, expression of *tfec* declines as iridoblasts become specified (Fig. 3.16; $tfec^{-/-}$ simulation). To experimentally test this prediction, expression of *tfec* itself in *tfec* mutant samples was investigated. This was possible due to the nature of the mutation, which did not lead to premature truncation of the transcript. At 24 hpf, 7/22 embryos (most likely mutants according to Mendelian ratios; one experimental repeat; 0.3 < p-value < 0.5) presented with lack of migrating iridoblasts and expansion of the *tfec* positive premigratory NCC domain, in a manner similar to that observed in $sox10^{t3}$ mutants (refer to section 3.2.2), suggesting specification defects in these embryos (Fig. 4.10 A,B).

By 2 dpf, a surprising phenotype was observed in 7/19 assessed embryos (one experimental repeat; 0.5 < p-value < 0.7; Fig. 4.11 E-H). *tfec* expressing cells were severely reduced in the RPE region and along the yolk sac stripe of these presumptive mutants. Cells were reduced but a subset was still present in positions of the developing LPs, while the swim bladder epithelium completely lacked *tfec* expression. Unexpectedly, *tfec* positive cells along the dorsal stripe were reduced by approximately 50% and cells of the ventral stripe by approximately 30% (Fig. 4.12). The intensity of staining in remaining cells along the dorsal stripe did not appear affected compared to

presumptive WT siblings, whereas most cells along the ventral stripe expressed *tfec* weakly (Fig. 4.11 E-H). Overall, these results suggested that *tfec* mutants harboured specification defects and highlighted a persistent population of cells distributed in an iridophore-like manner along the dorsal stripe, which did not express other iridophore markers, such as *ltk* and *pnp4a*.

In this context, *mitfa* expression dynamics were predicted to behave similarly to the WT, specifically there was an initial rise and a subsequent decline of Mitfa concentration (Fig. 3.16; *tfec*$^{-/-}$ simulation). To test the accuracy of this prediction, *mitfa* expression was examined in batches of *tfec* mutants and WT siblings. At 24 hpf, 5/25 assessed embryos (most likely mutants according to Mendelian ratios; one experimental repeat; $0.5 < p\text{-value} < 0.7$) displayed reduction of *mitfa* positive migrating cells and persistence of *mitfa* expression in the premigratory NC domain of the dorsal trunk (Fig. 4.10 C, D). This phenotype was attributed to specification defects and was linked to the reduced melanophore phenotype observed in live embryos at 30 hpf (refer to section 4.2.3). *mitfa* expressing cNCCs also appeared reduced, although this effect could be attributed to a subset of embryos being slightly younger than their siblings. At 2 dpf, *mitfa* expression appeared unaffected in melanocyte locations. 3/24 embryos, assessed in a single experimental repeat (most likely mutants according to Mendelian ratios; $0.1 < p\text{-value} < 0.2$), demonstrated an increase in *mitfa* positive cells along the length of the lateral stripe (data not shown), which correlated with the increase in melanocytes observed in live *tfec* mutant embryos (refer to section 4.2.3). Confirming either of those effects requires further experimental replicates and quantitation. These results indicate that initial upregulation of *mitfa* within the *tfec* positive population (Fig. 4.10 B) does occur, but it is not yet clear whether *mitfa* is subsequently downregulated from iridophore precursors. Further experiments at later stages are required to confirm whether model A predictions are accurate in this context.

According to model A, *pnp4a* expression initially becomes upregulated and subsequently declines with time upon loss of *tfec* function (Fig. 3.16; *tfec*$^{-/-}$ simulation). The predicted behaviour was supported by analysing the effects of knocking-down *tfec* using morpholino technology on *pnp4a* expression at 24 hpf (refer to section 3.2.6). Samples containing heterozygous and homozygous *tfec* mutant embryos as well as WT siblings were used to characterise these effects more robustly. Of the embryos subjected to *in situ* hybridisation, 5/23 (most likely mutants according to Mendelian ratios; one experimental repeat; $0.7 < p\text{-value} < 0.8$) showed prominent reduction in numbers of stained migrating cells and of cells located along the dorsal trunk (Fig. 4.10. E,F). These populations of cells were previously shown to be *mitfa*-dependent (refer to section 3.2.6).

Furthermore, cNCCs and cells overlaying the RPE, both regions that normally expressed *pnp4a,* appeared severely affected in presumptive *tfec* mutants.

By 2 dpf, a strong reduction in *pnp4a* positive iridophores was observed in 5/28 examined embryos (most likely mutants according to Mendelian ratios; one experimental repeat; 0.7 < p-value < 0.8; Fig. 4.11 I-L; Fig. 4.12). In affected embryos, faintly stained cells scored along the dorsal and ventral stripes did not appear to localise in iridophore positions but their distribution more closely resembled that of differentiated melanophores (data not shown). Confirmation of whether *mitfa*-dependent *pnp4a* upregulation takes place in *tfec* mutants at stages when differentiated cells are present is outstanding. These preliminary results indicated that model A accurately simulated *pnp4a* expression dynamics.

Finally, *sox10* expression was assessed in *tfec* mutants at 18 hpf (N = 21, one experimental repeat) and at 24 hpf (N = 25, one experimental repeat) in order to evaluate the prediction from model A, namely that expression of *sox10* gradually declines upon loss of *tfec* function, similarly to the WT (Fig. 3.16; *tfec*$^{-/-}$ simulation). Although the sample sizes were insufficient, in particular when subtle phenotypes were involved, a potential delay in migration of *sox10* positive cells was observed at 18 hpf (data not shown). The effect was quantified by scoring the number of streaks formed by ventrally migrating cells along the trunk. In presumptive WT and heterozygotes, 6 streaks were observed on average, but in 3/21 embryos (most likely mutants according to Mendelian ratios; one experimental repeat; 0.5 < p-value < 0.7), 1 streak on average was present in the anterior-most region of the trunk. In the affected embryos the observed streaks formed of very few cells compared to WT streaks (data not shown). Keeping in mind that the number of affected embryos was small (n = 3), hence this effect might be due to experimental artefacts, a t-test showed a statistically significant difference between the two observed phenotypes (Fig. 4.12 B). At 24 hpf, only 1/25 embryos (one experimental repeat) showed noticeably more prominent staining along the premigratory NC domain (unlikely a mutant according to Mendelian ratios; 0.01 < p-value < 0.05; data not shown). This result was considered a likely experimental artefact and indicated that expression of *sox10* in the premigratory NCC was not affected in *tfec* mutants at the examined stages, in line with model A predictions.

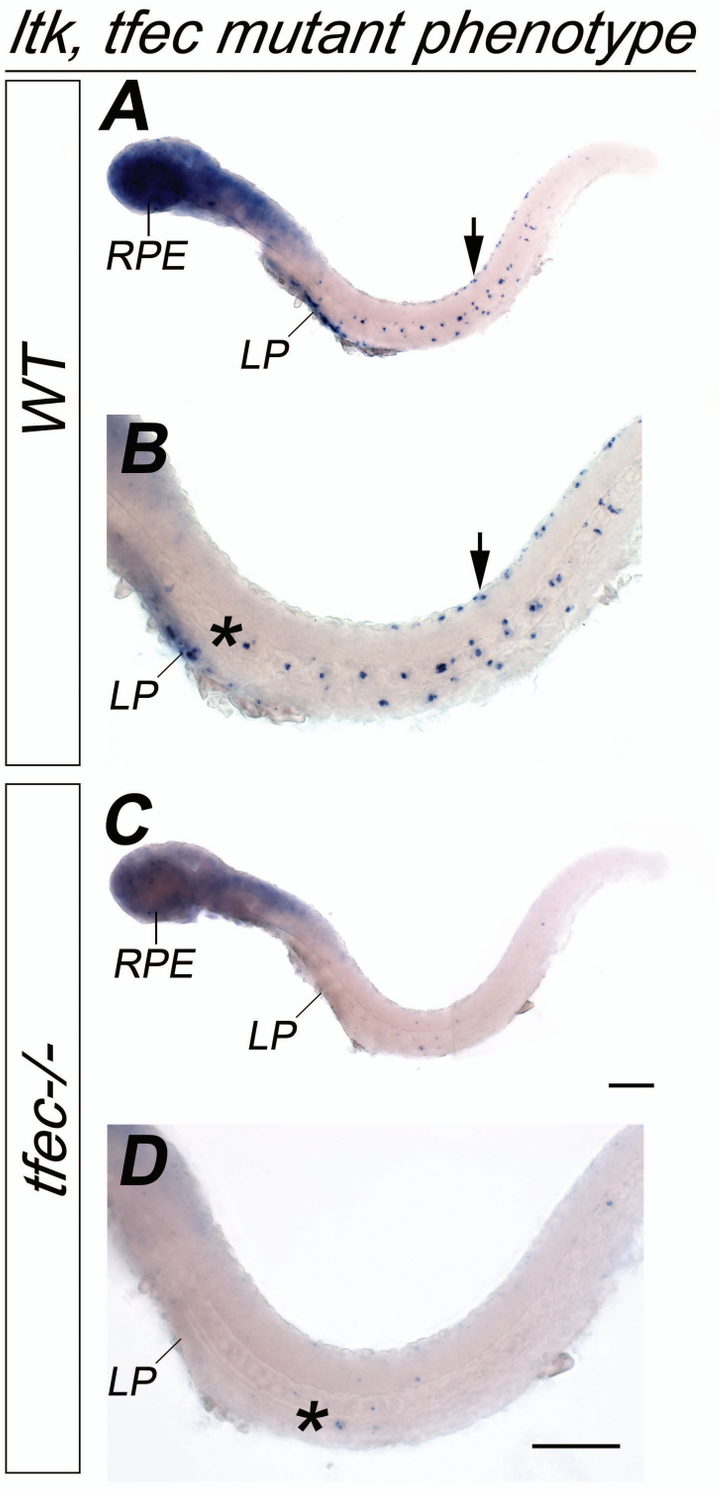


Figure 4.9. *ltk* expression is affected in *tfec* mutants during specification stages. *In situ* hybridisation data on presumptive WT (A, B) and *tfec* mutant (C, D) embryos at 30 hpf. In presumptive mutants, specified *ltk* positive iridoblasts are not detectable along the dorsal trunk (arrows) and on the medial migration pathway (asterisks). *ltk* is expressed in the vicinity of the RPE and in LP progenitors in the WT embryos, but not in affected siblings. RPE, retinal pigment epithelium; LP, lateral patches. Lateral views, oriented with the head to the left. Scale bars: 100 µm.

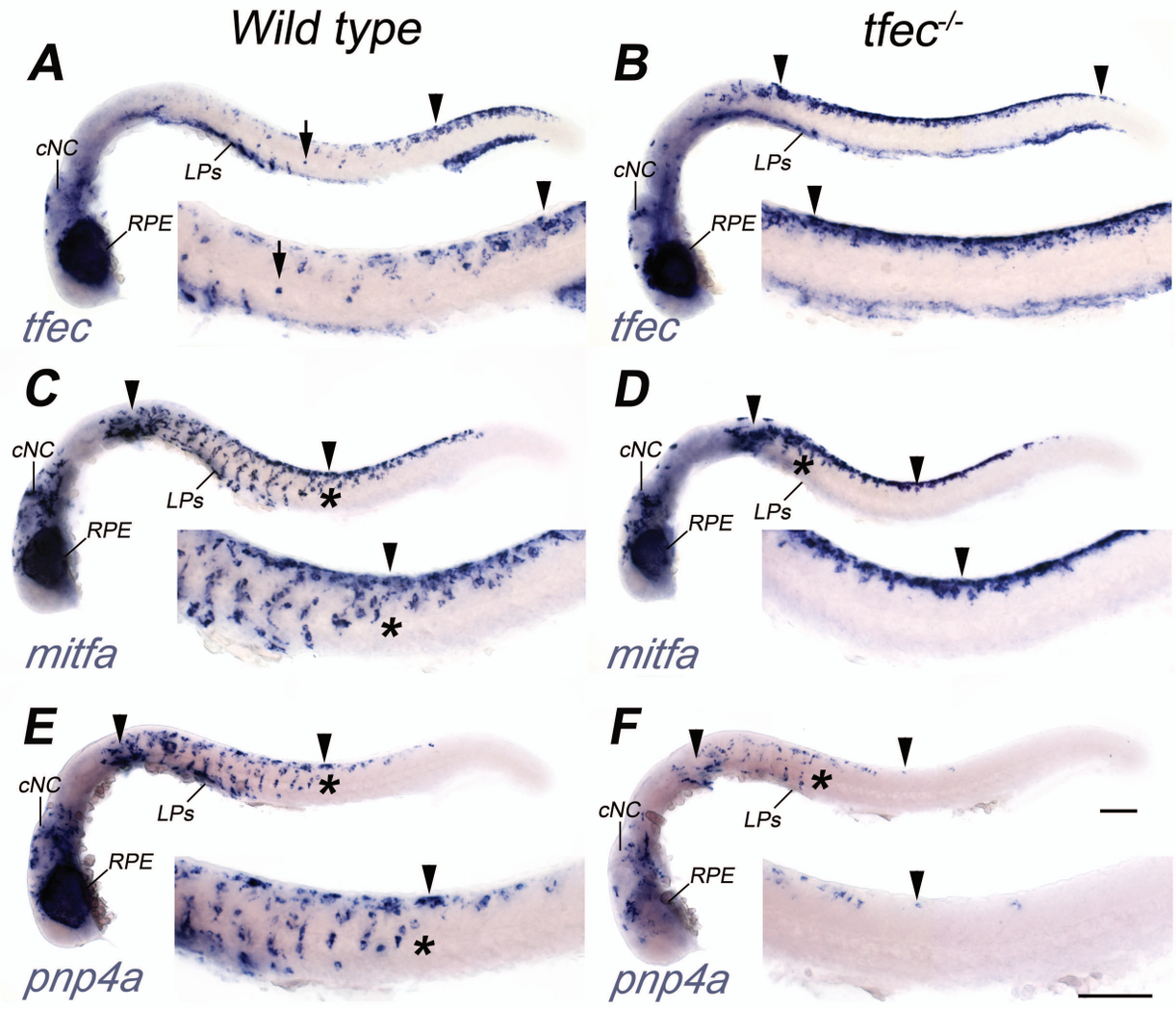


Figure 4.10. *tfec*, *mitfa* and *pnp4a* expression is affected in presumptive *tfec* mutants at 24 hpf. *In situ* hybridisation data on phenotypically WT (A, C, E) and presumptive *tfec* mutant (B, D, F) embryos at 24 hpf. (A, B) *tfec* is normally expressed in posteriorly located tNCCs, which appear increased and extending towards the anterior in presumptive mutants (arrowheads). Similarly, cNCCs fail to downregulate *tfec*. *tfec* positive migrating iridoblasts are missing in presumptive mutants (arrows). Expression in association with the RPE and LPs is not noticeably affected. (C, D) *mitfa* positive tNCCs (arrowheads), but not cNCCs, appear increased and a prominent lack of migrating (asterisks) and ventrally positioned melanoblasts is observed. Expression in the vicinity of the RPE is not obviously affected. (C, E) *pnp4a* is expressed in cNCCs, tNCCs (arrowheads), migrating cells (asterisks) and proximal to the RPE in WT embryos but expression in these domains is severely reduced in presumptive mutant siblings. LPs, lateral patches; RPE, retinal pigment epithelium; cNC, cranial neural crest. Lateral views, oriented with the head to the left. Scale bars: 100 µm.

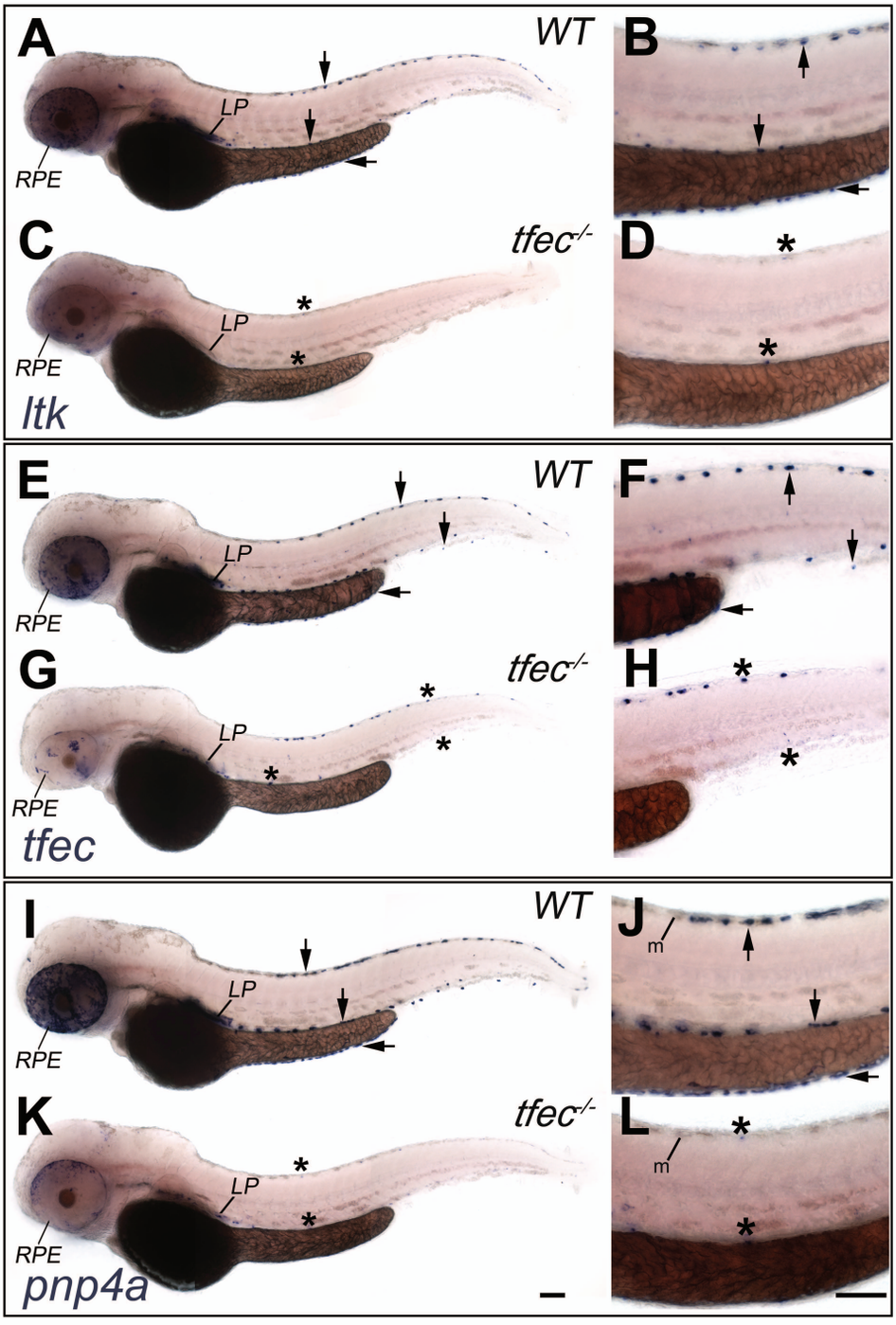


Figure 4.11. In presumptive *tfec* mutants at 2 dpf, *ltk* and *pnp4a* are no longer expressed in iridophore locations, but *tfec* positive cells are still observed. *In situ* hybridisation data on phenotypically WT (A, C, E, G, I, K) and *tfec* mutant (B, D, F, H, J, L) embryos at 2 dpf. In presumed mutants, differentiated iridophores expressing *ltk*, *tfec* and *pnp4a* along the dorsal (vertical arrows), ventral (vertical arrows) and yolk sac stripes (horizontal arrows), the LPs and overlaying the RPE are reduced. Very few escapers (asterisks) expressing *ltk* and *pnp4a* are present at iridophore positions (C, D, K, L), but an increased number of cells expressing *tfec* is detectable along the dorsal stripe (G, H). RPE, retinal pigment epithelium; LP, lateral patches; m, melanin (leftover after PTU treatment). Lateral views, oriented with the head to the left. Scale bars: 100 µm.

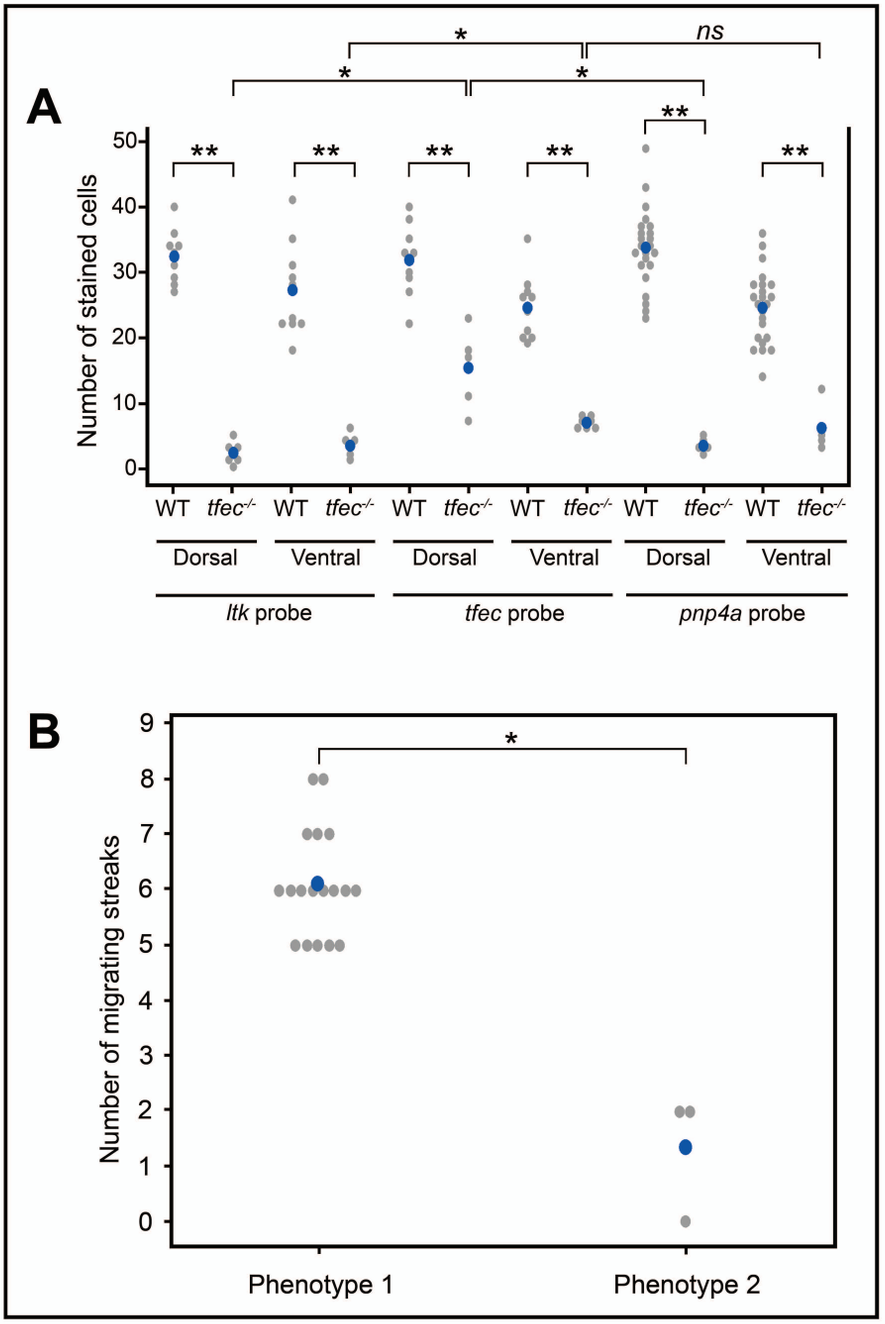


Figure 4.12. Scoring *tfec* mutant expression phenotypes. (A) Number of cells expressing *ltk*, *tfec* and *pnp4a* in phenotypically WT and presumptive *tfec* mutants at 2 dpf along the dorsal and ventral stripes. (B) Number of *sox10* positive migrating streaks at 18 hpf in presumed WT and mutant embryos. Statistical significance derived from two-sample t-tests: * p-value<0.05, ** p-value<0.001. ns, not significant.

Taken together, these preliminary data suggested that model A correctly predicted the early dynamics of all genes (between 18 hpf and 24 hpf). Specifically, *sox10*, *mitfa* and *tfec* appeared to be strongly expressed in the premigratory domain of the dorsal trunk, where the modelling begins. The predicted gradual elimination of Tfec between approximately 24 hpf and 40 hpf has yet to be experimentally tested. *ltk* was predicted to never become upregulated in this context, a hypothesis favoured by experiments to date. Finally, Pnp4a levels were predicted to rise prior to declining. In comparison to predictions in the other contexts (refer to sections 4.3.1, 4.3.2 and 4.3.3), Pnp4a peaked at relatively low levels, even though the parameter values were not altered across simulations. Although experiments to date indicated that *pnp4a* expression in *tfec* mutants was likely to be weaker than in WT siblings (Fig. 4.10 E,F), support from quantitative data is required to test this prediction.

4.3.5. Evaluating the loss of *mitfa* function context

In line with experimental evidence, simulating loss of *mitfa* function predicted no differences from WT expression dynamics, with the exception of *pnp4a* dynamics, which presented with loss of the *mitfa*-dependent 'peak' before reaching the plateau phase. No conclusions could be drawn from this observation, as the model was not designed to predict quantitative changes. Moreover, this WT 'peak' of Pnp4a concentration could be attributed to parameter choice. As was shown in section 3.2.6, *pnp4a* expression was significantly affected in *mitfa*^{w2} mutants at 24 hpf, and was at later stages restored in presumptive iridoblasts, the cell population which this GRN aimed to model. This prediction was considered to be biologically inaccurate, as it failed to reflect initial loss of *pnp4a* expression in *mitfa* mutants.

Transient delay of *ltk* expression in *mitfa*^{w2} mutants had been previously demonstrated in the research group (S. Lopes, PhD thesis). However, this observation could not be predicted in this model, as parameter choice could bias the modelling output and possibly obscure relevant changes.

Regarding *tfec* expression dynamics in *mitfa* mutants, samples comprising WT embryos, heterozygous carriers and homozygous *mitfa*^{w2} mutants were subjected to *in situ* hybridisation to detect *tfec* transcript. At 18 hpf, when *tfec* was normally expressed in the cNCCs and tNCCs, no observable alterations were present across two experimental replicates (N ≈ 50). This was consistent with data showing *mitfa* expression in only a small subset of cells of the anterior dorsal trunk at this stage (Fig. 3.15 G,H; Fig. 4.6 A). At 24 hpf, 21/62 embryos (most likely mutants according to Mendelian ratios; 3 independent experimental repeats; $0.1 < p\text{-value} < 0.2$) showed a very strong phenotype of *tfec* expression maintenance along the premigratory NCC domain, while *tfec* was being downregulated in siblings showing the WT expression pattern (Fig. 4.13 A,B). *tfec*

was also expressed more prominently than in phenotypically WT embryos in cNCCs and in NCCs proximal to the hindbrain/trunk boundary (Fig. 4.13 A,B). Both of these regions were normally labelled by *mitfa* and by *pnp4a* expression (Fig. 4.10 C; Fig. 3.11 A,B). Finally, when compared to WT siblings, more intense expression was observed in the developing LPs and increased numbers of ventrally migrating cells were detectable in presumptive mutants. Expression in the vicinity of the RPE appeared unaffected.

At 30 hpf an analogous phenotype was observed in 13/39 embryos (most likely mutants according to Mendelian ratios; 2 independent experimental repeats; $0.2 < \text{p-value} < 0.3$; Fig. 4.13 C,D). Judging by more intense and compact staining, the number of cells expressing *tfec* along the dorsal trunk appeared increased compared to phenotypically WT siblings. Additionally, an increased number of ventrally migrating cells was detectable and expression in the developing LPs was more pronounced. Similar phenotypes were observed amongst N = 39 embryos at 36 hpf, assessed in 2 independent experiments ($0.5 < \text{p-value} < 0.7$; Fig. 4.13 E,F).

Although the number of cells expressing *tfec* was detectably increased in *mitfa*^{w2} mutants, in agreement with published data showing increased numbers of mature iridophores in these mutants (J A Lister et al. 1999), it could not be concluded that the dynamics in developing iridoblasts were affected without performing quantitative experiments. Therefore, the current understanding is that model A accurately predicts expression of both *tfec* and *ltk*, but not of *pnp4a* upon loss of *mitfa* function.

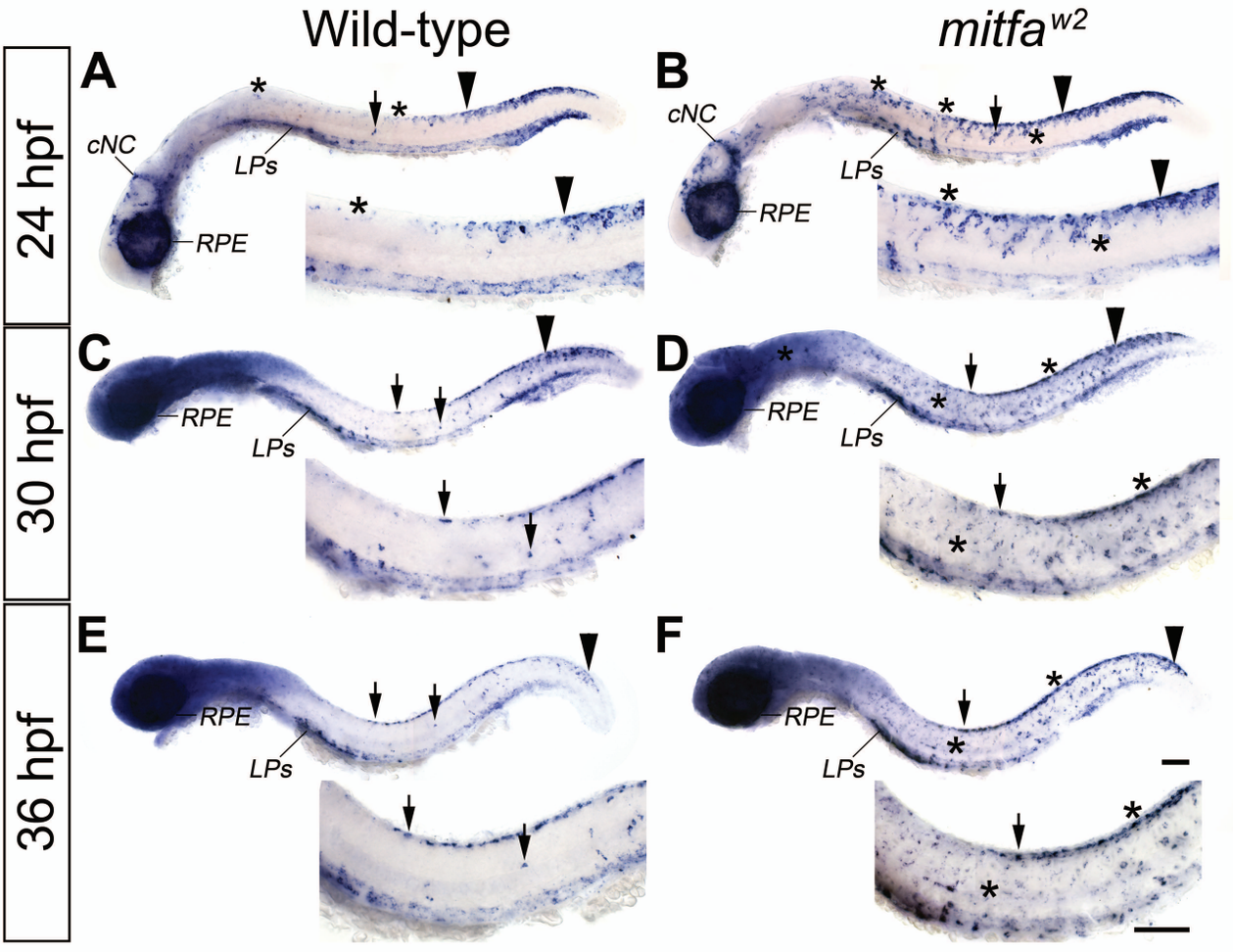


Figure 4.13. *tfec* expression fails to be downregulated in NCC derivatives in *mitfa*^{w2} mutants. *In situ* hybridisation data on phenotypically WT (A, C, E) and presumptive *mitfa*^{w2} mutant (B, D, F) embryos at 24 hpf (A, B), 30 hpf (C, D) and 36 hpf (E, F). *tfec* is normally expressed in cNCCs, premigratory tNCCs (arrowheads), specified migrating iridoblasts (arrows), over the RPE and on the LPs. In affected embryos, tfec is expressed in WT locations but, instead of being downregulated at 24 hpf (A, asterisks), expression persists in premigratory, migrating and cranial NC derivatives (B, D, F, asterisks). RPE, retinal pigment epithelium; LP, lateral patches; cNC, cranial neural crest. Lateral views, oriented with the head to the left. Scale bars: 100 μm.

4.4. Model B: *sox10* is essential for *pnp4a* upregulation

4.4.1. Hypothesis and mathematical modelling

Model A inaccurately predicted *pnp4a* expression dynamics upon loss of *sox10* function. Although *tfec* expression was shown to persist in the premigratory NC domain of *sox10* mutants (refer to section 3.2.2), *pnp4a* expression was only activated in very few cells and reached very low levels by 30 hpf (refer to section 4.3.2). It was thus concluded that *sox10* directly or indirectly mediated *tfec*-dependent *pnp4a* expression.

Moreover, it was demonstrated that at 24 hpf, *mitfa* upregulated *pnp4a* expression (refer to section 3.2.6). *pnp4a* was expressed at 24 hpf in a manner reminiscent of *mitfa*, but was gradually downregulated in *mitfa*-expressing melanoblasts as specification progressed. Ultimately, *pnp4a* was solely expressed as an iridophore marker (Fig. 3.11). These observations were interpreted taking into account published data (Greenhill et al. 2011), which established the dynamics of *sox10* expression in differentiated melanophores. It was shown that at 28 hpf, *sox10* was expressed in 100% of melanophores, but that percentage was gradually reduced until 48 hpf, when 0% of melanophores expressed *sox10*. Model B was formulated on the hypothesis that gradual downregulation of *sox10* in melanophores (Greenhill et al. 2011) was responsible for the observed downregulation of *pnp4a* in this cell type.

In summary, the proposed model B (Fig. 4.14) featured a requirement for *sox10* function for both *tfec*-dependent and *mitfa*-dependent *pnp4a* activation. In the corresponding diagram (Fig. 4.14), this effect is represented using black arrows from *sox10* towards the relevant edges of the network, indicating that *sox10* function is essential for these interactions to occur.

To theoretically test the formulated hypotheses, model B was mathematically described using a system of ODEs. Up to this point, mathematical modelling was carried out assuming presence of different binding sites for each TF on a promoter (non-competitive activation) and implementing the OR gate, meaning that binding of any one activating TF alone can trigger transcription from that promoter. To incorporate the notion of *sox10* requirement, AND gates were introduced in the modelling process (Greenhill et al. 2011), meaning that two or more TFs were necessary for transcription to take place. As described by Greenhill *et al*., when implementing OR gates, the probability of gene G to be transcribed after binding of transcriptional activators A and/or B is calculated as follows:

$$p(G^*) = p(A, \sim B) + p(\sim A, B) + p(A, B) \qquad (26)$$

whereas if both A and B are required for transcriptional activation, the probability is the following:

$$p(G^*) = p(A, B) \tag{27}$$

Using equations (26) and (27), the probability of *pnp4a* activation by *sox10* and *tfec*, *sox10* and *mitfa* or by all three TFs can be calculated as follows:

$$p(P^*) = p(S, T, \sim M) + p(S, \sim T, M) + p(S, T, M) \tag{28}$$

From (28), following guidelines provided by Greenhill *et al.*, the following ODE is derived to describe the rate of change of Pnp4a concentration over time:

$$\frac{d[P^P]}{dt} = g_S \times [r_S^P] \times ([r_T^P] - [r_M^P] \times [r_T^P] + [r_M^P]) - d_P \times [P^P] \tag{29}$$

where

$$[r_M^P] = \frac{[M^P]}{Kd_M^P + [M^P]} \tag{30}$$

$$[r_T^P] = \frac{[T^P]}{Kd_T^P + [T^P]} \tag{31}$$

$$[r_S^P] = \frac{[S^P]}{Kd_S^P + [S^P]} \tag{32}$$

The equations describing the rates of change of the remaining genes as well as the parameter constants used for model B were as described in chapter 3 (refer to section 3.3.3).

4.4.2. Theoretical evaluation

To examine the predictions of model B theoretically, MATLAB software was employed to numerically solve the newly derived system of ODEs. The output (Fig. 4.15) was used to theoretically test the accuracy of the model. Qualitatively, predictions of model B concerning *ltk* and *tfec* mutant contexts were the same as in model A, where they were shown to be consistent with biological observations to date (refer to section 3.3.4). These predictions will not be further discussed in this section. Interestingly, in model B predicted gene expression dynamics upon loss of *sox10* function were consistent with experimental observations. Specifically, *pnp4a* expression was predicted to remain undetectable throughout development, as was experimentally demonstrated (refer to section 4.3.2), whereas expression of both *ltk* and *tfec* was sustained (refer to sections 3.2.2 and 3.2.3). However, the WT and loss of *mitfa* function simulations were not biologically accurate. *pnp4a* was shown to be expressed in differentiated iridophores (refer to section 3.2.6), therefore its predicted downregulation in the WT context by model B was inaccurate. This behaviour arose due to the manually implemented *sox10* downregulation in the lineage. Finally, loss of *mitfa* function resulted in *pnp4a* downregulation, for the same reasons as in the WT simulation (Fig. 4.15; *mitfa*$^{-/-}$ simulation).

Overall, theoretical testing of model B highlighted the fact that if *pnp4a* expression was indeed *sox10*-dependent, as was experimentally suggested, then downregulation of *sox10* in the iridophore lineage was not appropriate. Given that downregulation of *sox10* was an assumption based on data from the literature and computational tests, and based on model B predictions was likely invalid, expression of the gene was tested experimentally. *sox10* dynamics in the iridophore lineage were investigated using *in situ* hybridisation as well as the more sensitive RNAscope technique.

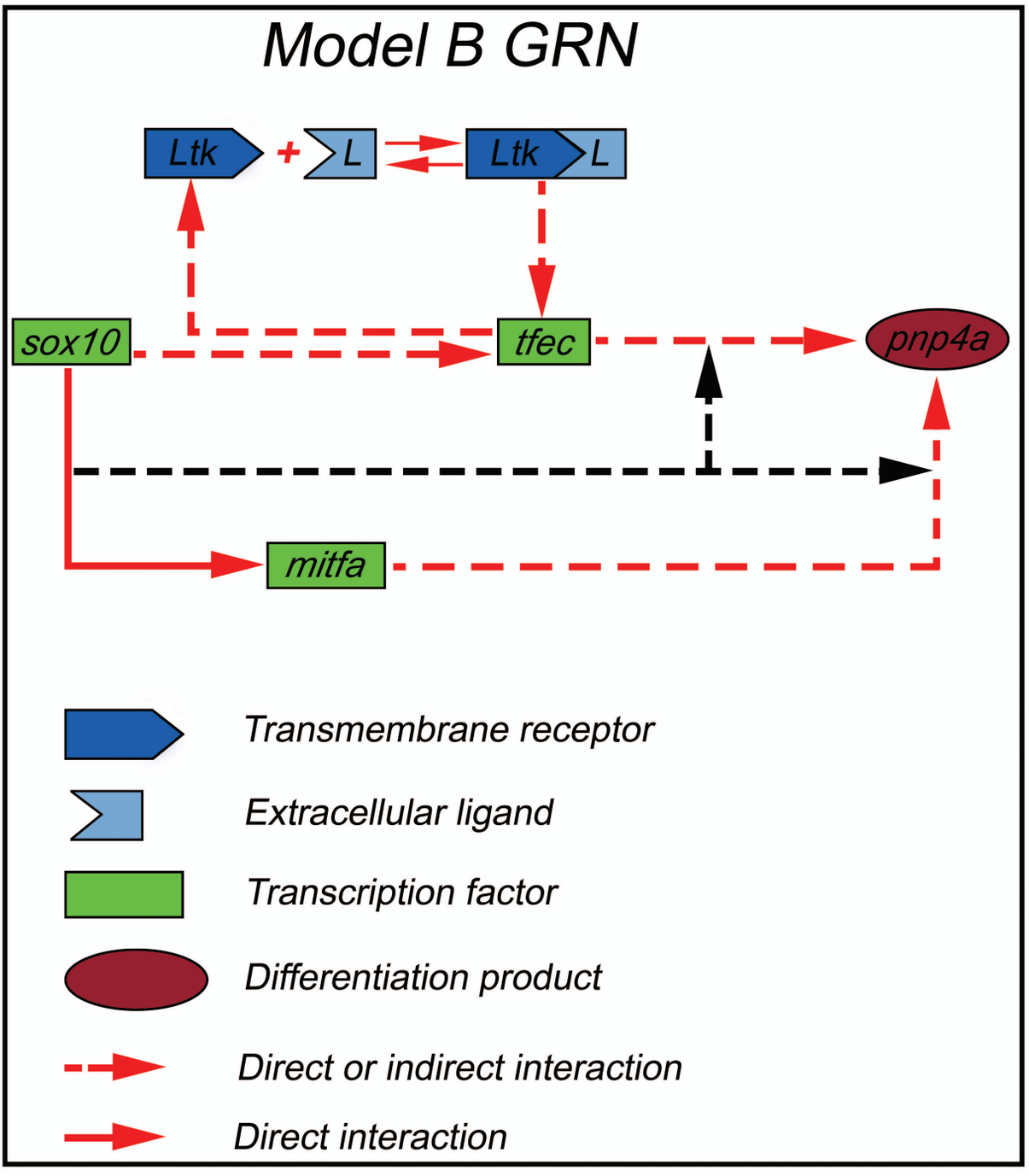


Figure 4.14. Model B. The key indicates symbols used for different types of molecules and interactions. All nodes are represented using gene names, except for Ltk, which is shown as the RTK binding to its ligand (L) to form a complex that activates gene expression. Model B was derived from model A by adding dashed arrows (black) from *sox10* towards the arrows between *tfec* and *pnp4a* and between *mitfa* and *pnp4a*. These indicated that *sox10* cooperates with both *tfec* and *mitfa* to upregulate *pnp4a*.

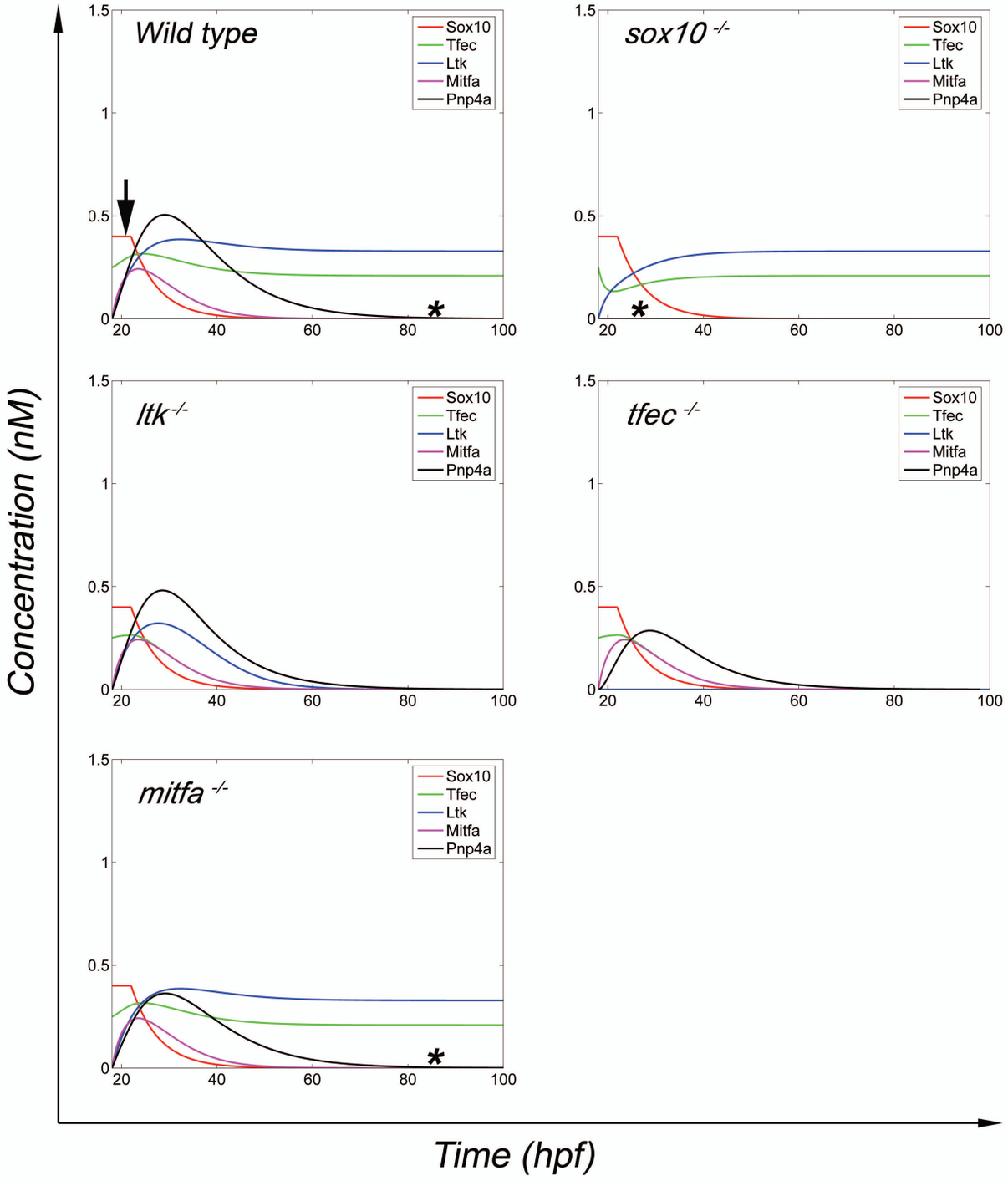


Figure 4.15. Predicted gene expression dynamics in model B. Model B inaccurately predicts downregulation of *pnp4a* in the WT and loss of *mitfa* function contexts (asterisks). In the *sox10* loss of function context the prediction is biologically relevant (asterisk). This downregulation was a result of *sox10* decline in the iridophore lineage over time (WT, arrow). The remaining predictions are accurate based on experimental data to date.

4.4.3. *sox10* is maintained throughout iridophore development

The first step towards examining *sox10* expression dynamics in the iridophore lineage was by performing conventional *in situ* hybridisation to determine whether *sox10* transcript was present in mature iridophores. Two live WT embryos were imaged at 2 dpf under incident light, and independently subjected to *in situ* hybridisation using a DIG-labelled RNA probe complementary to the *sox10* mRNA sequence. *sox10* expression was detected in known locations, including Schwann cells both of the posterior lateral line and associated with spinal nerve axons, oligodendrocytes and the otic epithelium (Fig. 4.16 B,D). Surprisingly, prominent *sox10* expression was also observed in all positions occupied by iridophores, possibly with the exception of cells associated with the RPE. Specifically, *sox10* transcript was present in iridophores of the dorsal, ventral and yolk sac stripes, as well as in the LPs (Fig. 4.16 A-D). To reinforce this finding, *in situ* hybridisation was performed on embryos carrying the *ltk*^{ty82} mutant allele at 60 hpf. Iridophore-like *sox10* expression was eliminated in a subset of the embryos, presumably corresponding to *ltk*^{ty82} homozygous mutants (Fig. 4. 16 E,F).

Following the establishment of *sox10* expression in mature iridophores, the gene's expression dynamics were assessed over the course of specification of the iridophore lineage. Co-expression studies were conducted using RNAscope, with the aim of determining whether expression of the iridoblast marker *ltk* co-localised with *sox10* expression throughout iridophore development.

RNAscope was performed on WT embryos (WIK background) fixed at 24 hpf, 30 hpf and 48 hpf. At all stages *ltk* positive cells were invariably *sox10* positive (Fig. 4.17). At 24 hpf, *sox10* expression could be seen in migrating NCC derivatives which strongly expressed *ltk* (Fig. 4.17 A-D, arrowheads). Low levels of *ltk* expression were detected in migrating (Fig. 4.17 A-D, asterisks), as well as in premigratory (data not shown), NCCs positive for *sox10* expression. This pattern is not fully visible when performing conventional *in situ* hybridisation (Fig. 3.5) (Lopes et al. 2008) due to the relatively low signal detection threshold of the technique.

At 30 hpf, all *ltk* positive NCC derivatives detected along the dorsal trunk and the medial migration pathway expressed *sox10* (Fig. 4.17 E-H). Finally, at 48 hpf, N = 112 *ltk* positive cells residing along the dorsal trunk in iridophore locations were scored for *sox10* expression (data not shown). Of these cells, 93.8% expressed *sox10*. It was therefore established that *sox10* expression persisted throughout iridophore development.

Conclusively, these results demonstrate that *sox10* expression persists in the majority, if not all, of specified iridoblasts and mature iridophores in zebrafish embryos.

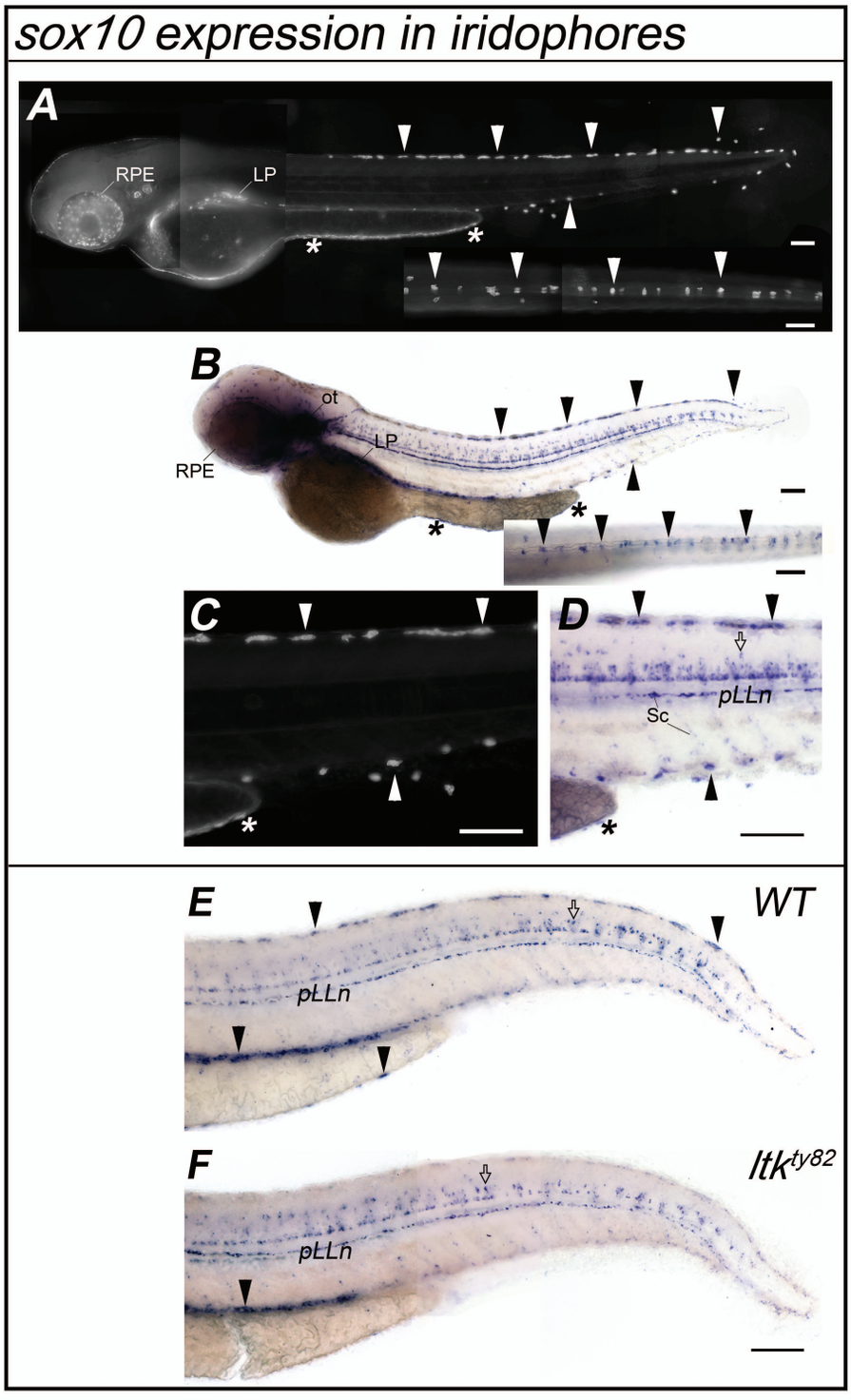


Figure 4.16. *sox10* is expressed in mature iridophores. A PTU-treated WT embryo imaged at 2 dpf under incident light (A, C) subjected to *in situ* hybridisation to detect *sox10* transcript (B, D). *sox10* expression labels expected cell types (Sc, pLLn, oligodendrocytes indicated by empty arrow) and iridophore locations along the dorsal and ventral stripes (arrowheads), the yolk sac stripe (asterisks) and on the LPs. At 60 hpf, presumptive *ltk*^{ty82} mutants (F), compared to WT siblings (E), presented with loss of *sox10* expression in the majority of iridophore locations (arrowheads). RPE, retinal pigment epithelium; LP, lateral patches; pLLn, posterior lateral line nerve; Sc, Schwann cells; ot, tic epithelium. Lateral views, oriented with the head to the left. Scale bars: 100 µm.

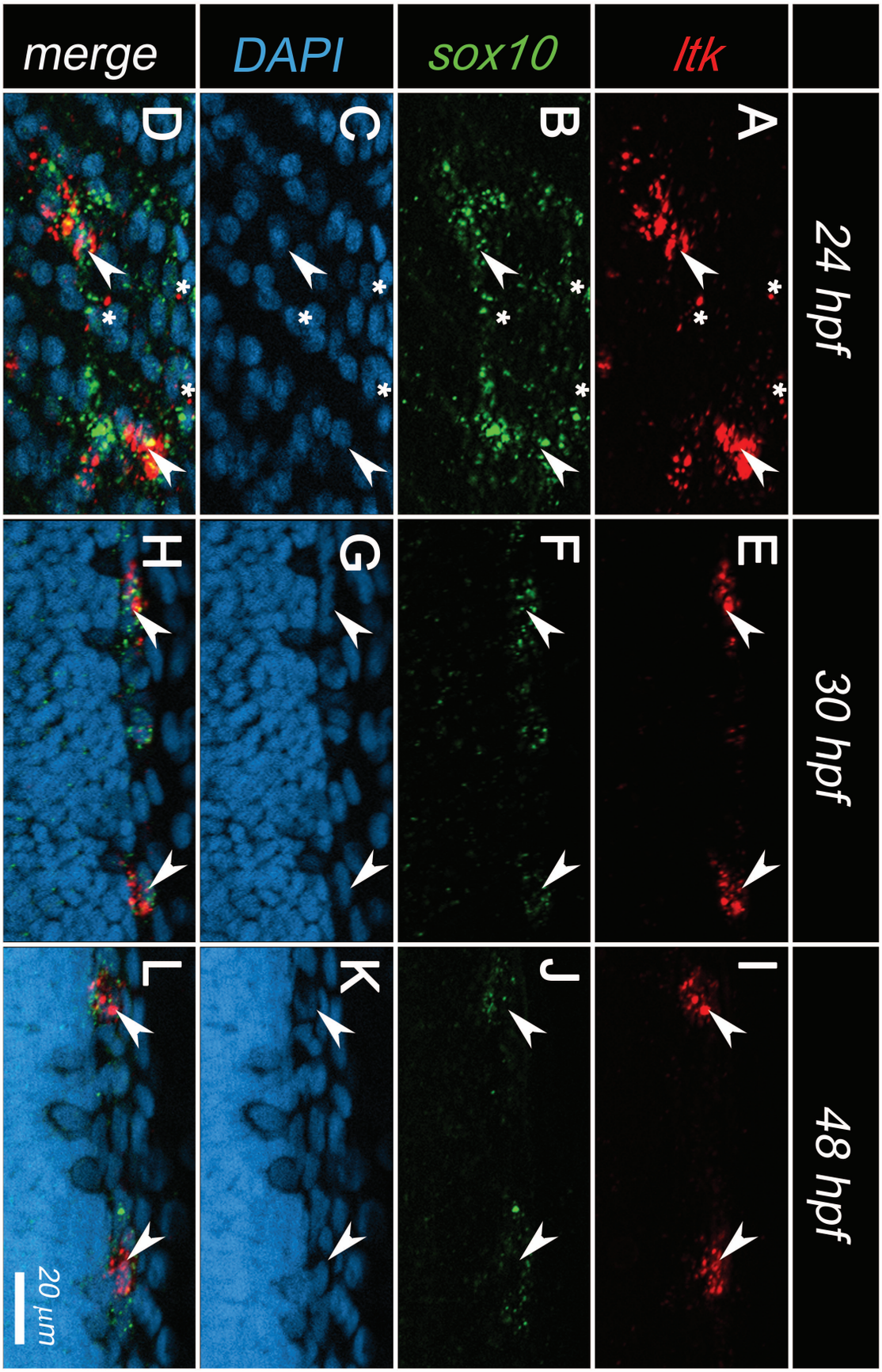


Fig. 4.17

Figure 4.17. *sox10* is expressed in *ltk*-positive iridoblasts and iridophores. Co-labelling of *ltk* (red) and *sox10* (green) transcripts in WT embryos at 24 hpf, 30 hpf and 48 hpf using RNAscope. At 24 hpf (A-D), *sox10* is expressed in dorsally located *ltk*+ presumptive iridoblasts, which are entering the medial migratory pathway (arrowheads). Asterisks indicate low level of *ltk* expression in other medially migrating NCCs. At 30 hpf (E-H), *ltk*-expressing iridoblasts along the dorsal trunk (arrowheads) are positive for *sox10* expression. At 48 hpf (I-L) cells identified at iridophore positions along the dorsal stripe (arrowheads) co-express *ltk* and *sox10*. For each stage, presented panels are derived from a single Z-slice of the corresponding Z-stack file. Lateral views, oriented with the head to the left. Scale bar: 20 μm.

4.4.4. Experimental testing

Experimental testing of model B hypotheses, namely that *sox10* was required for both *tfec*-dependent and *mitfa*-dependent *pnp4a* upregulation, was attempted by conducting mRNA co-overexpression studies (refer to section 2.7). Preceding overexpression data by Dr Tatiana Subkhankulova, a postdoctoral fellow in the group, had shown that overexpressing each of *sox10*, *tfec* or *mitfa* individually was insufficient to drive ectopic *pnp4a* expression at 6 hpf (unpublished data; results not shown). In this study, *sox10*/*mitfa* or *sox10*/*tfec* transcripts were simultaneously overexpressed and downstream effects on gene expression were assessed at 6 hpf and at 10 hpf using q-RT PCR, with the aim of identifying whether *sox10* was required as a cofactor for *pnp4a* upregulation.

For each co-overexpression experiment, the following construct combinations were injected in WT (WIK background) embryos at the single-cell stage:

1) GFP mRNA (reference sample for q-RT PCR).
2) Mutant transcripts of both TFs. The following constructs were used: $sox10^{m618}$, $mitfa^{b692}$ and GFP mRNA instead of a *tfec* mutant allele, due to lack of *tfec* mutant sequences at the time. For example, $sox10^{m618}$/$mitfa^{b692}$ co-injection acted as a negative control for *sox10*/*mitfa* co-expression.
3) Functional version of one factor, combined with the mutant version of another. For example, $sox10^{WT}$/$mitfa^{b692}$ or $sox10^{m618}$/$mitfa^{WT}$.
4) Functional versions of both factors of interest. For example, $sox10^{WT}$/$mitfa^{WT}$

Approximately 30 pg of each transcript were deposited in the yolk sac of each embryo. To account for variability of injected mRNA levels between individual embryos, total RNA was extracted from pools of 8 embryos per sample. cDNA was synthesised and q-RT PCR was performed. Relative gene expression levels were determined using the ΔΔCT method (refer to section 2.2.3.3). Two-tailed t-test was used to indicate statistical significance (p-value < 0.05) of the relative gene expression mean fold changes in samples injected with either one or two functional transcripts compared to samples injected with two non-functional transcripts (negative control). This was because overexpressing non-functional mRNA molecules was considered a more appropriate negative control than overexpressing GFP. For each condition, two biological replicates were processed.

To determine whether *sox10*, together with *mitfa*, could drive *pnp4a* expression, the following transcript combinations were injected: GFP alone, $sox10^{m618}$/$mitfa^{b692}$, $sox10^{WT}$/ $mitfa^{b692}$, $sox10^{m618}$/ $mitfa^{WT}$ and $sox10^{WT}$/ $mitfa^{WT}$. The effect of overexpressing $mitfa^{WT}$ transcript was validated by confirming that the melanocyte marker *dct*, a known target of *mitfa*, was upregulated by 6 hpf upon injection of the relevant construct (Fig. 4.18 A).

The effect of overexpressing the *sox10* construct was validated by confirming its ability to ectopically trigger *mitfa* expression at 10 hpf (Fig. 4.18 D). In this experiment, *pnp4a* expression was not significantly upregulated in embryos at any of the assessed stages (Fig. 4.18 A,B). Moreover, a statistically significant 2-fold increase in *tfec* expression was observed upon injection of *sox10* alone, which did not manifest when WT *mitfa* was injected (Fig. 4.18 B).

These results indicated that the combination of Sox10 and Mitfa was insufficient to induce *pnp4a* expression either at 6 hpf or at 10 hpf, suggesting that further cofactors present in NCCs are necessary for activating the gene in NC derivatives. Furthermore, the observed 2-fold increase in *tfec* expression upon *sox10* mRNA injection supported the notion that *sox10* upregulates *tfec* in iridoblasts (refer to section 3.2.2), although efficient activation of *tfec* may require additional cofactors. The fact that co-overexpressing *sox10* and *mitfa* did not result in *tfec* upregulation was in line with experimental evidence arguing that *mitfa* represses *tfec* expression (Fig. 4.13), although further biological replicates are required to confirm this result.

To determine whether co-injecting *sox10* and *tfec* had an effect on *pnp4a* transcription, the following transcript combinations were injected: GFP alone, *sox10*^{m618}*/GFP, sox10*^{WT}*/GFP, sox10*^{m618}*/tfec* and *sox10*^{WT}*/tfec.* Successful injection of *tfec* transcript was evaluated at 6 hpf (Fig. 4.18 C), however no targets responding to *tfec* overexpression have been reported thus far, therefore the functionality of the injected construct could not be confirmed. Results from this experiment did not demonstrate *pnp4a* upregulation with any combination of injected transcripts either at 6 hpf or at 10 hpf. However, statistically significant repression of *pnp4a* was suggested at 10 hpf upon *tfec* overexpression both in the presence and absence of functional *sox10* (Fig. 4.18 C,D). This result was considered a likely experimental artefact, as it did not fit with loss of function data that suggested *tfec* is important for *pnp4a* upregulation (Fig. 4.10), and awaits confirmation.

In summary, these experiments are to be repeated, but strongly suggested that *pnp4a* upregulation required cofactors or signals yet to be identified, which are present in the NCC microenvironment, but not in embryos between 6 hpf and 10 hpf. Even though co-overexpression studies thus far failed to support the hypotheses based on which model B was generated, strong evidence derived from loss of function studies, indicated that *sox10* is essential for *pnp4a* upregulation (refer to sections 4.3.2 and 4.4.1), despite the exact molecular mechanisms remaining elusive. These results are discussed further in section 4.7.

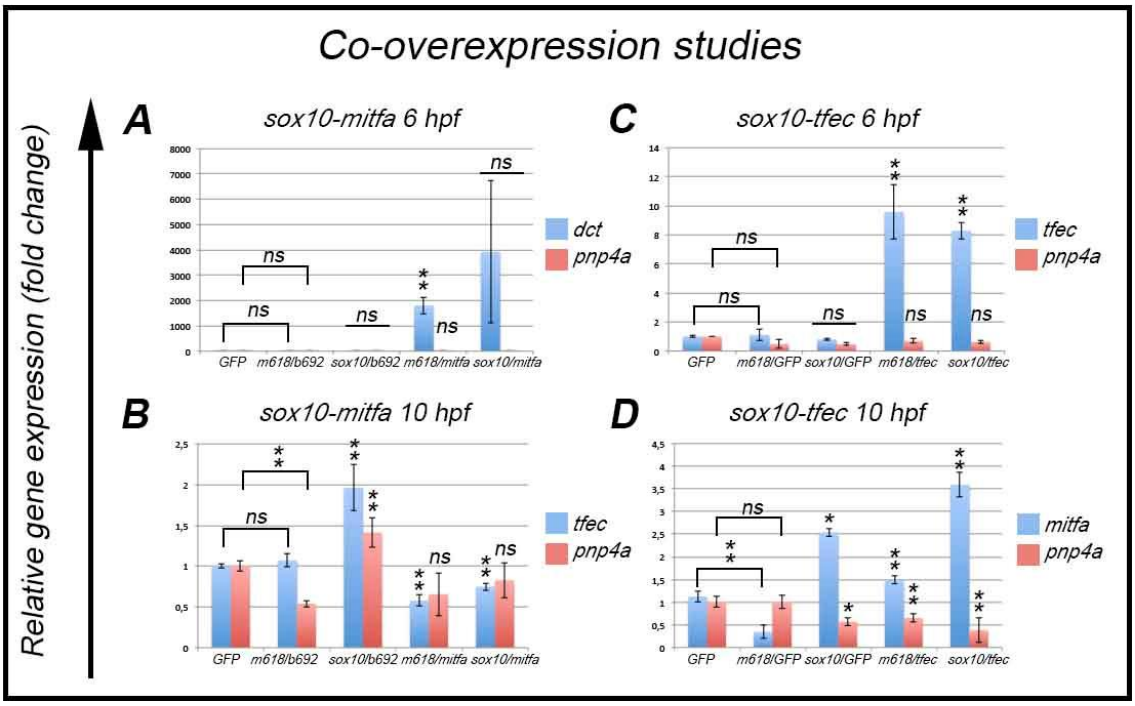


Figure 4.18. *pnp4a* was not ectopically expressed following co-overexpression of *sox10/tfec* or of *sox10/mitfa*. q-RT PCR results showing the mean gene expression fold change between GFP injected samples and co-overexpression samples. Error bars indicate standard error of the mean based on two experimental replicates per condition. Two-tailed t-test was used to establish statistical significance of the mean fold change when one or two functional transcripts were injected, compared to the mean fold change when two null transcripts were co-injected. Successful injection of *tfec* was validated based on robust increase of *tfec* expression levels (C). Successful upregulation of *dct* following WT *mitfa* transcript injection (A) and of *mitfa* following WT *sox10* transcript (D) were confirmed. *sox10*, together with *mitfa* failed to upregulate *pnp4a* expression at 6 hpf (A) and at 10 hpf (B). *tfec* was weakly but significantly upregulated by WT *sox10* injection at 10 hpf (B). *sox10* together with *tfec* also failed to upregulate *pnp4a* expression at 6 hpf (C) and at 10 hpf (D). Instead, a statistically significant reduction in relative *pnp4a* expression levels was observed at 10 hpf (D). ** p-value < 0.05; * p-value < 0.08; ns, not significant.

4.5. Model C: *sox10* is upregulated in an iridophore-specific manner

4.5.1. Hypothesis and mathematical modelling

Theoretically testing model B indicated that the assumption regarding *sox10* downregulation in the iridophore lineage was likely inaccurate. Indeed, *sox10* maintenance throughout iridophore development was experimentally demonstrated. It was further shown that *sox10* was expressed in the majority, if not in all, of mature iridophores. For *sox10* to be maintained, it followed that it needed to be regulated in an iridophore-specific manner.

A generic way of achieving *sox10* expression in the iridophore lineage involved addition of an iridogenic factor I in the GRN, which would directly or indirectly drive *sox10* expression (Fig. 4.19 A). The unknown nature, dynamics and upstream effectors or downstream targets of factor I would render fitting it in the network and experimentally testing associated predictions impossible. Keeping this generic model in mind, the network was simplified by hypothesising that *sox10* was upregulated in a *tfec*-dependent manner, thus implying the formation of a positive feedback loop between *sox10* and *tfec* (Fig. 4.19 B). This was considered an appropriate assumption since *tfec* was being regarded as the iridophore master regulator based on its expression pattern and mutant phenotype. Advantageously, *tfec*-dependent *sox10* upregulation can be tested experimentally using *tfec* mutants and overexpression studies. Moreover, additional assumptions in regards to parameter choice, which would have arisen if the generic model were to be implemented, were avoided. The resulting model C is depicted in figure 4.19 C.

tfec-dependent upregulation of *sox10* was incorporated in the mathematical model by deriving the following equation to describe the rate of change of Sox10 concentration:

$$\frac{d[S^P]}{dt} = g_S \times \left[r_T^S\right] - d_S \times [S^P] \qquad (33)$$

where

$$\left[r_T^S\right] = \frac{[T^P]}{Kd_T^S + [T^P]} \qquad (34)$$

The dissociation constant value for this equation was the same as the one used previously to describe *tfec*-dependent *ltk* and *pnp4a* upregulation:

$$Kd_T^S = 0.3\ nM$$

Similarly the rates of *sox10* mRNA production and protein degradation were as before:

$$g_S = 0.2\ nM/h$$

$$d_S = 0.175\ \frac{1}{h}$$

The above equation replaced the Heaviside function previously used to regulate Sox10 input. The equations describing the rates of change of other genes were the same as in model B.

4.5.2. Theoretical evaluation

To theoretically test the consistency of model C predictions with experimental observations, the system of ODEs incorporating *tfec*-dependent *sox10* upregulation was numerically solved using MATLAB. Importantly, computational outputs indicated that *pnp4a* dynamics, both in the WT and in the loss of *sox10* function context (Fig. 4.20), were consistent with experimental observations. Specifically, *pnp4a* remained upregulated in the WT, consistent with data presented in figure 3.11, while expression failed to be activated in the loss of *sox10* function context, in line with experimental observations (Fig. 4.7). Therefore, maintenance of *sox10* expression in the iridophore lineage improved the predictive accuracy of the model with regards to *pnp4a* regulation. Nevertheless, examination of model C predictions in different contexts revealed two considerable inaccuracies.

In the WT simulation, consistently with known biology, all iridophore marker genes, namely *tfec*, *ltk*, *pnp4a* and *sox10*, remained upregulated throughout development of the lineage (Fig. 4.20; WT simulation). However, continuous activation of the melanophore gene, *mitfa*, was also observed, due to it being directly targeted by *sox10* (Elworthy et al. 2003). This prediction was deemed inaccurate, as *mitfa* is not expressed in mature iridophores (Fig. 4.6; Lister et al. 1999). Not surprisingly, continuous *mitfa* expression in mature iridophores was also predicted in both the *ltk* and *mitfa* loss of function contexts, where functional *sox10* was maintained (Fig. 4.20). It was, therefore, concluded that a repressor for *mitfa* is required to downregulate the gene over the course of iridophore development.

All predictions in the loss of *sox10* function context reflected experimental observations (Fig. 4.20; *sox10*$^{-/-}$ simulation). Specifically, *sox10* expression was sustained over time, as has been previously established (Dutton et al. 2001), similarly to *tfec* and *ltk* expression, both of which were demonstrated to remain upregulated in the premigratory NC domain (refer to sections 3.2.2 and 3.2.3). *mitfa* expression was never initiated in this context, in line with previous reports (Elworthy et al. 2003).

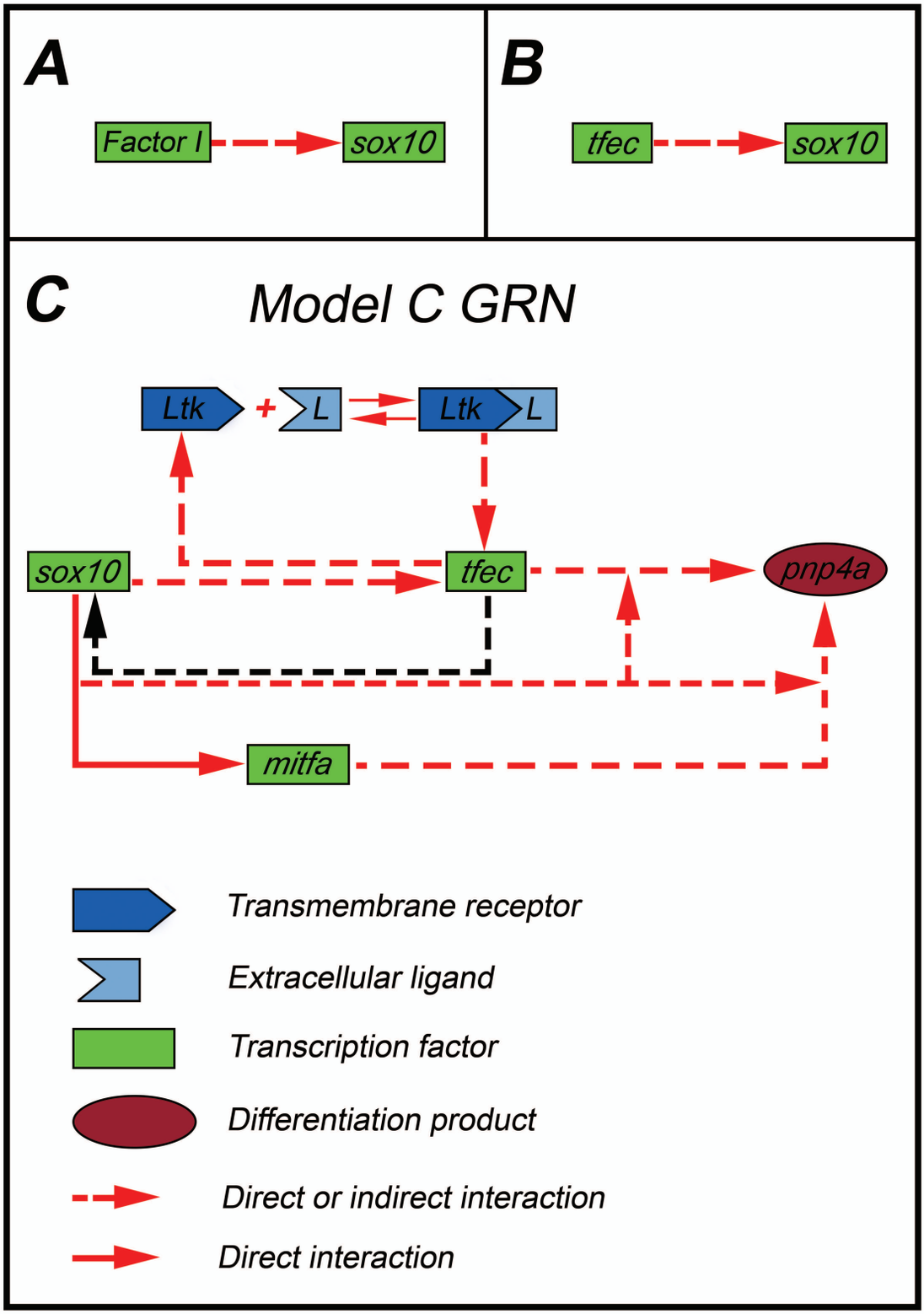


Figure 4.19. Model C. The key indicates symbols used for different types of molecules and interactions. All nodes are represented using gene names, except for Ltk, which is shown as the RTK binding to its ligand (L) to form a complex that activates gene expression. Model C was derived from model B by adding the black arrow from *tfec* towards *sox10*, indicating the formation of a positive feedback loop between the two genes.

Intriguingly, upon loss of *ltk* function, model C predictions presented with significant inaccuracies (Fig. 4.20; *ltk*$^{-/-}$ simulation). In particular, gene expression dynamics appeared qualitatively unchanged when compared to the WT simulation. Considering the established striking effects of loss of *ltk* function on the expression of both *tfec* and *pnp4a* (refer to sections 3.2.5 and 4.3.3, respectively), predictions in this context were considered biologically inaccurate. Moreover, in the case of *ltk* mutants, Lopes *et al*. have shown that *ltk* expression emerged in NCCs, but became rapidly downregulated at later developmental stages (Lopes et al. 2008). Therefore, the prediction that loss of *ltk* function does not affect Ltk concentration in the iridophore lineage was not in line with known biology. The highlighted loss of *ltk* dependency in model C was due to the *tfec*/*sox10* positive feedback loop being sufficient for maintaining *tfec* and, hence, *pnp4a* expression in the absence of functional *ltk*. Theoretical testing of alternative models was performed in an attempt to restore *ltk* dependency in model C (refer to section 4.5.3).

In model C, predictions upon loss of *tfec* function were qualitatively consistent with experimentally established gene regulatory dynamics, similarly to previous models (refer to section 4.3.4). Finally, simulated loss of *mitfa* function did not negatively impact on expression of iridogenic genes, thus the predictions were considered accurate with the exception of maintenance of *mitfa* expression throughout development.

Taken together, these data indicated two inaccuracies in model C. Firstly, upregulation of the melanophore-specific gene, *mitfa*, in the iridophore lineage, as a result of being directly targeted by Sox10. Secondly, lack of dependency on *ltk* function for the development of the iridophore lineage. This outcome was due to the positive feedback loop between *sox10* and *tfec* maintaining expression of both genes and of *pnp4a* even in the absence of *ltk*. It should be noted that even if the more generic model were adopted, continuous *sox10* expression in the lineage would have resulted in *mitfa* and *tfec* upregulation. *tfec* and *sox10* would have, subsequently, cooperatively upregulated *pnp4a* in the lineage even in the absence of *ltk*.

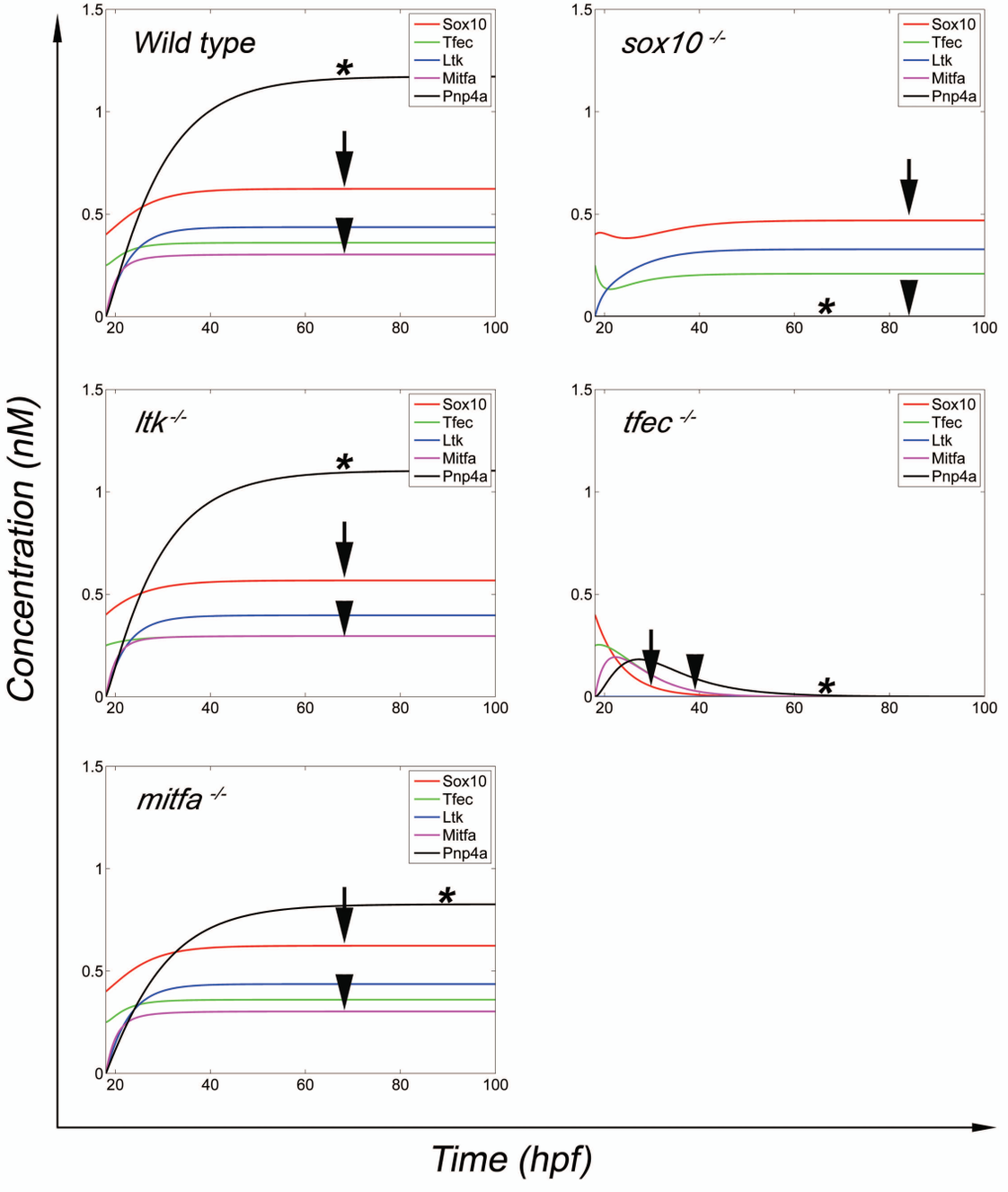


Figure 4.20. Predicted gene expression dynamics in model C. Model C inaccurately predicts continuous *mitfa* expression in the WT, *ltk* mutant and *mitfa* mutant context (arrowheads) as a result of *sox10* upregulation in the iridophore lineage (arrows). The predictions upon loss of *ltk* function are also biologically inaccurate, as iridogenic genes like *pnp4a* (asterisks) remain upregulated throughout development. According to experimental evidence to date, all other predictions are biologically accurate, including elimination of *pnp4a* upon loss of *sox10* function (asterisk).

4.5.3. Theoretically testing alternative models – *ltk* dependency

In order to address the inconsistencies between model C predictions in the loss of *ltk* function context and experimental data, a set of testable hypotheses were considered. These hypotheses were first evaluated theoretically, using the mathematical model and, subsequently, experimentally tested (refer to section 4.5.4). The following three alternatives for model C were theoretically tested:

C.1) *ltk* is essential for *sox10* to upregulate *tfec* in the iridoblast (Fig. 4.21 A).
C.2) *ltk* is essential for *tfec* to upregulate *sox10* in the iridoblast (Fig. 4.21 B).
C.3) *ltk* is essential for *tfec* to upregulate *sox10*, but also upregulates *tfec* in a non *sox10*-mediated fashion (Fig. 4.21 C).

In the first two cases, the effect of *ltk* on *tfec* was elicited through *sox10*. In the third case, *ltk* was assumed to upregulate *tfec* expression both in a *sox10*-dependent and in a *sox10*-independent fashion.

For these tests, the objective was to assess whether any model would produce biologically appropriate predictions in the WT and in the loss of *sox10* and loss of *ltk* function contexts, since the models' predictions have thus far been proven biologically inaccurate in these particular contexts. When assessing simulation outputs, the ectopic maintenance of *mitfa* expression in the WT context was not considered as this was an independent issue which needed to be tackled separately.

Firstly, expression dynamics were assessed if *ltk* was essential for *sox10* to activate *tfec* expression in an iridophore-specific manner (model C.1; Fig. 4.21 A). To mathematically model this feature of the network, the following equation was used to describe the rate of change of Tfec concentration over time:

$$\frac{d[T^P]}{dt} = g_T \times [r_S^T] \times [r_L^T] - d_T \times [T^P] \qquad (35)$$

where

$$[r_S^T] = \frac{[S^P]}{Kd_S^T + [S^P]} \qquad (36)$$

$$[r_L^T] = \frac{[L^P]}{Kd_L^T + [L^P]} \qquad (37)$$

and

$$Kd_S^T = 0.2 \times 10^{-9}\, M$$

$$Kd_L^T = 0.1 \times 10^{-9}\, M$$

$$g_T = 0.2\, nM/h$$

$$d_T = 0.5\ h^{-1}$$

When implementing this change to the model, loss of *ltk* function predictions were consistent with biological observations, indicating that the role of *ltk* was restored in the iridophore lineage. However, *sox10* mutants were no longer predicted to maintain *tfec* and *ltk* expression (Fig. 4.21 A), thus this hypothesis was disregarded.

Next, the assumption that *ltk* was essential for the positive feedback loop because it enabled *tfec*-dependent *sox10* upregulation was tested (model C.2; Fig. 4.21 B). This relationship was modelled mathematically by allowing *tfec* to only be upregulated by *sox10* (equation 1) and by implementing an AND gate dictating that both *ltk* and *tfec* were essential for *sox10* upregulation (equation 2) as follows:

$$\frac{d[T^P]}{dt} = g_T \times [r_S^T] - d_T \times [T^P] \qquad (38)$$

$$\frac{d[S^P]}{dt} = g_S \times \left[r_T^S\right] \times \left[r_L^S\right] - d_S \times [S^P] \qquad (39)$$

where

$$[r_S^T] = \frac{[S^P]}{Kd_S^T + [S^P]} \qquad (40)$$

$$[r_L^T] = \frac{[L^P]}{Kd_L^T + [L^P]} \qquad (41)$$

$$\left[r_L^S\right] = \frac{[L^P]}{Kd_L^S + [L^P]} \qquad (42)$$

All parameters remained as in model C.1 and the additional dissociation constant $Kd_L^S = 0.1 \times 10^{-9}\, M$ was added. The output indicated restoration of *ltk* dependency during iridophore development, while in the WT context all iridogenic genes behaved according to experimental observations. However, model C.1 predicted that both *tfec* and *ltk* would be downregulated during iridophore development in the loss of *sox10* function context (Fig. 4.21 B). This prediction is biologically inaccurate since both genes have been demonstrated to remain upregulated in trapped premigratory precursors in *sox10* mutants (Fig 3.3, 3.5 and 3.6).

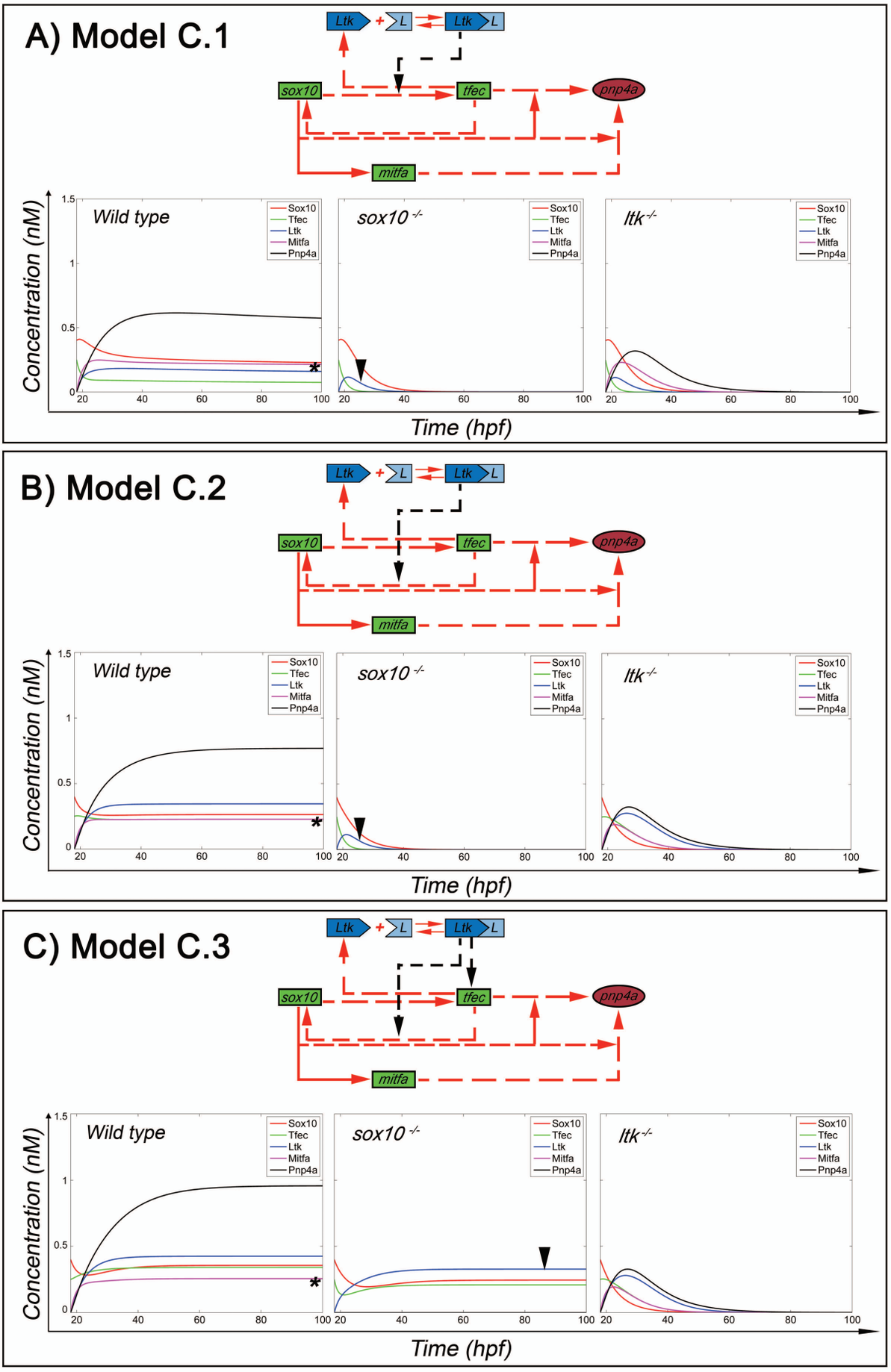

A) Model C.1
Ltk
+
L
Ltk
L
sox10
tfec
pnp4a
mitfa
Wild type
sox10 -/-
ltk -/-
Sox10
Tfec
Ltk
Mitfa
Pnp4a
Concentration (nM)
Time (hpf)
B) Model C.2
C) Model C.3

Fig. 4.21

Figure 4.21. Theoretically testing alternative models to restore *ltk* dependency. Models C.1, C.2 and C.3 were sequentially proposed to restore the requirement for Ltk signalling during iridophore development. Under C.1 (A), Ltk signalling is hypothesised to exert its effects on *tfec* by mediating *sox10*-dependent *tfec* upregulation. Under C.2 (B), Ltk signalling is presumed essential for *tfec*-dependent *sox10* upregulation to take place. Under C.3 (C), Ltk signalling is hypothesised to be essential for the establishment of the feedback loop between *sox10* and *tfec*, but also to upregulate *tfec* in a *sox10*-independent manner. C.1 and C.2 inaccurately predict *tfec* and *ltk* downregulation in a *sox10* mutant context (A, B, arrowheads). C.3 accurately predicts maintenance of both *tfec* and *ltk* in the *sox10* mutant context (C, arrowhead) and downregulation of iridogenic genes in the *ltk* mutant context. In the WT simulation, all three models' predictions are accurate, except for continuous *mitfa* upregulation (A, B, C, asterisks).

Models C.1 and C.2 both predicted downregulation of *tfec* and *ltk* expression upon loss of *sox10* function. This was because in both cases the *ltk*/*tfec* positive feedback loop only operated via *sox10*. Whilst simulating models C.1 (Fig. 4.21 A) and C.2 (Fig. 4.21 B) produced similar predictions, dynamics associated with loss of *ltk* function more closely resembled the observed expression patterns in C.2 predictions. Specifically, *tfec* expression in C.1 was shown to plunge immediately at the start of modelling, while in C.2 expression formed a peak, correlating to the gene being initially expressed in unspecified NCCs and specified iridoblasts (refer to section 3.2.5). Therefore, C.2 was modified to derive model C.3, according to which *ltk* regulated *tfec* both through mediating *tfec*-dependent *sox10* upregulation via an AND gate, and by exerting effects on *tfec* in a *sox10*-independent manner (Fig. 4.21 C).

For mathematical modelling of model C.3, the following set of equations was derived to describe *tfec* (equation 43) and *sox10* (equation 44) expression dynamics:

$$\frac{d[T^P]}{dt} = g_T \times ([r_S^T] \times (1 - [r_L^T]) + [r_L^T]) - d_T \times [T^P] \quad (43)$$

$$\frac{d[S^P]}{dt} = g_S \times [r_T^S] \times [r_L^S] - d_S \times [S^P] \quad (44)$$

where

$$[r_S^T] = \frac{[S^P]}{Kd_S^T + [S^P]} \quad (45)$$

$$[r_L^T] = \frac{[L^P]}{Kd_L^T + [L^P]} \quad (46)$$

$$[r_L^S] = \frac{[L^P]}{Kd_L^S + [L^P]} \quad (47)$$

All parameter constants remained the same as in models C.1 and C.2.

Simulating model C.3 suggested that all protein concentrations behaved as expected according to experimental observations in the WT context, with the exception of Mitfa. Upon loss of *sox10* function, *tfec*, *ltk* and *sox10* were predicted to remain expressed, in line with data presented in sections 3.2.2 and 3.2.3 and in published literature (Dutton et al. 2001), respectively. Finally, *ltk* dependency was restored: predictions in the loss of *ltk* function context were consistent with experimental observations (refer to sections 3.2.5

and 4.3). Expression dynamics upon loss of *tfec* function also reflected the known biology using this model (data not shown).

In conclusion, given the current connectivity within the GRN, incorporating the hypothesis that Ltk signalling is responsible for upregulating *tfec* both in a *sox10*-dependent and in a *sox10*-independent manner generated biologically realistic predictions of gene expression dynamics.

4.5.4. Experimental testing

Theoretical testing using mathematical modelling suggested that the proposed model C.3 was able to generate predictions consistent with biological observations. Experimental testing was attempted in order to confirm that, as suggested by the modelling output, the interactions proposed in model C.3 were more biologically accurate than those in C.1 or in C.2, and thus that C.3 should be used as the basis of further iterations. For testing, co-overexpression studies were performed as described in section 4.4.4.

Experimental confirmation of the following requirements would support model C.1:

a. Overexpressing *sox10* does not induce *tfec* expression, except when co-overexpressed with *ltk.*
b. Overexpressing *tfec* alone induces ectopic expression of both *sox10* and *ltk.*
c. Overexpressing *ltk* does not ectopically activate *tfec*, except when co-overexpressed with *sox10.*
d. Overexpressing *ltk* alone does not activate *sox10* expression.

Moreover, validity of model C.2 required the following:

a. Overexpressing *sox10* induces *tfec* expression.
b. Overexpressing *tfec* alone induces ectopic expression of *ltk* and requires *ltk* function to subsequently activate *sox10*.
c. Overexpressing *ltk* does not ectopically activate *sox10,* except when co-overexpressed with *tfec*.
d. Overexpressing *ltk* alone does not activate *tfec* expression.

Finally, validity of model C.3 required the following:

a. Overexpressing *sox10* induces *tfec* expression.
b. Overexpressing *tfec* alone induces ectopic expression of *ltk* and requires *ltk* function to subsequently activate *sox10.*
c. Overexpressing *ltk* induces *tfec* expression, and then cooperates with *tfec* to induce *sox10* expression.

As before, for each co-overexpression assay the following combinations of transcripts were co-injected:

1) GFP (q-RT PCR reference sample).
2) Equal amounts of non-functional transcripts. *sox10*^{m618} transcript was injected as a negative control for *sox10* overexpression. The dead kinase version (*npm-ltk*(DK)) of the nucleophosmin *ltk* (*npm-ltk*) construct was injected as a negative control for *ltk* overexpression, which was achieved by injecting *npm-ltk* transcript, coding for a constitutively active form of Ltk (M. Nikaido, unpublished data). Both the *npm-ltk* and the *npm-ltk*(DK) constructs were previously generated in the group. The fusion protein coded for by *npm-ltk* transcript is able to, in the absence of a ligand, initiate the signalling pathways normally activated by endogenous ligand-bound Ltk. In the *npm-ltk*(DK) construct, a single amino acid substitution within the Ltk kinase domain renders the RTK inactive. Similarly to the experiments described in section 4.4.4, GFP was used as a negative control for *tfec* overexpression.
3) Equal amounts of one functional transcript and one mutant transcript.
4) Equal amounts of two functional transcripts.

Both the *npm-ltk* and the *npm-ltk(*DK*)* construct encoded truncated versions of the *ltk* transcript, starting from exon 16. Primers used for detecting *ltk* transcript via q-RT PCR were designed by Dr Tatiana Subkhankulova across the exon 8/exon 9 boundary (forward) and exon 9 (reverse). Therefore, ectopic activation of the complete endogenous transcript, but not of the injected truncated transcript was readily detectable.

Firstly, to investigate condition (a) for all 3 models, the following constructs were co-injected in WT siblings at the single-cell stage: GFP alone, *sox10*^{m618}*/npm-ltk(*DK*)*, *sox10*^{WT}*/npm-ltk(*DK*)*, *sox10*^{m618}*/npm-ltk* and *sox10*^{WT}*/npm-ltk*. For each combination, two biological replicates were processed and each replicate represented 8 injected siblings. For this experiment, embryos were only assessed at 6 hpf (Fig. 4.22 A). As expected, no significant increase in *tfec* expression levels was detectable between GFP-injected embryos and embryos injected with *sox10*^{m618}*/npm-ltk(*DK*)*. Consistently with experiments described in section 4.4.4, there was a statistically significant 2-fold increase in *tfec* transcript levels upon overexpression of WT *sox10*, but not increase was noted in *sox10*^{m618}*/npm-ltk(*DK*)* injected samples. However, neither *npm-ltk* nor *sox10*/*ltk* overexpression produced a statistically significant increase in relative *tfec* expression levels. These results suggested either that *sox10* alone could regulate *tfec* expression, an effect repressed through constitutively activation of Ltk signalling, or that activation of *tfec* expression at 6 hpf following injection of *sox10*/*npm-ltk(*DK*)* was an experimental artefact. Performing this experiment at 10 hpf is likely to provide a better understanding

of the overexpression consequences, since in other experiments gene expression changes were clearer at this time point (Fig. 4.22 B,C). Importantly, these results argued against validity of model C.1, since *sox10* and *ltk* co-overexpression did not appear to cooperatively activate *tfec*.

Subsequently, requirements (b), (c) and (d) were investigated. Regarding condition (b), it should be noted that *tfec* was expected to be able to upregulate *sox10* in each of the models C.1, C.2 and C.3, either independently of *ltk* (model C.1) or via Ltk signalling, following *tfec*-dependent induction of endogenous *ltk* (models C.2 and C.3), which was assumed to occur in all three models. Therefore, successful upregulation of *sox10* upon injection of *tfec* alone could not be definitively interpreted as ltk-independent, *tfec*-mediated upregulation, because endogenous *ltk* activation has likely occurred.

To investigate (b), (c) and (d), the following constructs were co-injected: GFP alone, *npm-ltk(*DK*)*/GFP, *npm-ltk*/GFP, *tfec*/*npm-ltk(*DK*)* and *npm-ltk*/*tfec*. As before, two biological replicates were processed for each combination and each replicate represented 8 injected siblings. Analyses at 6 hpf indicated no statistically significant changes in relative *sox10* or *tfec* expression levels between *npm-ltk(*DK*)*/GFP injected embryos and embryos injected with *npm-ltk*/GFP, *tfec/npm-ltk(*DK*)* or *npm-ltk/tfec* (Fig. 4.22 B). However, by 10 hpf, 10-fold higher *sox10* expression and 5-fold higher *ltk* expression were observed in *tfec/npm-ltk(*DK*)* injected embryos compared to GFP-injected embryos (Fig. 4.22 C). This suggested that *tfec* alone was able to activate endogenous *ltk* expression, as well as *sox10* expression, thus reinforcing condition (b) for models C.1, C.2 and C.3. It should be noted that for results of this assay at 10 hpf, GFP injected embryos were used as negative controls for statistical testing, due to unexpectedly high *sox10* expression in *npm-ltk(*DK*)*/GFP samples. This was likely due to embryos in those samples being slightly older than 10 hpf, thus starting to express endogenous *sox10*.

Interestingly, co-injection of *npm-ltk* and GFP transcripts did not result in ectopic upregulation of *sox10* or of endogenous *ltk* at 10 hpf (Fig. 4.22 C). According to requirements (c) and (d), these results are consistent with models C.1 and C.2 being valid, as opposed to model C.3, under which *ltk* overexpression is able to induce ectopic *tfec* expression. As stated before, model C.1 was excluded on the basis of apparent lack of cooperative activation between *sox10* and *ltk* towards inducing *tfec* expression (Fig. 4.22 A).

Moreover, co-injecting *tfec* and *npm-ltk* transcripts resulted in a 5-fold mean increase in *sox10* and *ltk* expression levels compared to negative control levels. Due to large error bars, the increase in relative *ltk* expression was not statistically significant using a two-tailed t-test (p-value = 0.07). These results resembled those obtained following injection

of *tfec*/*npm-ltk(*DK*)*, thus suggesting that *tfec* was responsible for the observed effects, and that a role for *ltk* is unlikely. Therefore, models C.2 and C.3, according to which Ltk signalling mediates *tfec*-dependent *sox10* activation, were contradicted.

Conclusively, theoretically testing alternative modifications of model C predicted that model C.3, but neither C.1 nor C.2, reflected observations derived from loss of function studies. However, contradictory evidence was obtained by preliminary co-overexpression tests that supported the proposed model C.2. Due to this controversy, and the fact that overexpression results require more experimental replicates to verify any drawn conclusions, none of the models C.1, C.2 or C.3 were incorporated in the GRN at this stage. The originally proposed model C (Fig. 4.19 C) was thus used for further iterations, while the unresolved predicted loss of *ltk* dependency remains an important issue to be addressed.

Finally, activation of the iridogenic marker *pnp4a* either at 6 hpf or at 10 hpf following co-overexpression of *tfec* and *ltk* was examined. Surprisingly, *pnp4a* upregulation was not observed in any of the samples, which may be due to lack of important cofactors that are expressed in the NC but not in embryos at 6 hpf and 10 hpf. Surprisingly, statistically significant downregulation of *pnp4a* was observed in *tfec*-injected samples. This result remains to be confirmed with further experimental replicates.

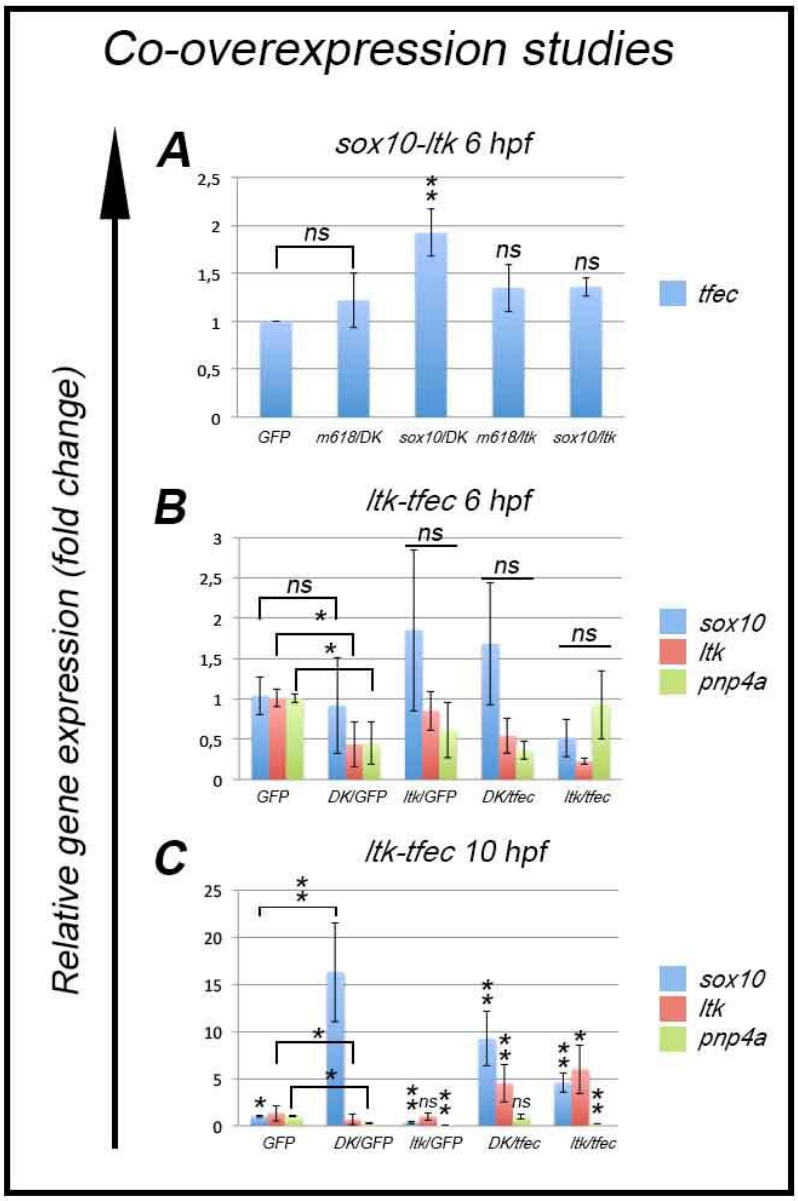


Figure 4.22. *tfec* overexpression results in ectopic upregulation of *sox10* and *ltk*, but not of *pnp4a*. q-RT PCR results showing mean gene expression fold change between GFP injected samples and co-overexpression samples. Error bars indicate standard error of the mean based on two experimental replicates per condition. Two-tailed t-test was used to establish statistical significance of the mean fold change when one or two functional transcripts were injected, compared to the mean fold change when two null transcripts were co-injected (A, B) or when GFP was injected alone (C). Overexpression of WT *sox10* results in weak but significant *tfec* upregulation (A). Co-overexpression of *tfec* and *ltk* does not produce statistically significant results at 6 hpf, although endogenous *ltk* and *pnp4a* expression appears potentially downregulated in negative controls, compared to the reference sample (B), but WT *tfec* upregulates *sox10* and *ltk* at 10 hpf (C). *pnp4a* expression was not upregulated, on the contrary statistically significant repression was observed upon injection of WT *tfec* transcript (C). ** p-value < 0.05; * p-value < 0.08; ns, not significant.

4.6. Theoretical testing of alternative models for *mitfa* repression

The role of *ltk* in the *sox10*/*tfec* feedback loop still lacked appropriate experimental support. To avoid adding unsupported features to the GRN, model C alternatives (C.1, C.2 and C.3) were rejected. Because WT *mitfa* upregulation in model C was independent to the observed loss of *ltk* dependency, model C (Fig. 4.19) was modified to derive alternative GRNs that predicted decline of *mitfa* expression in the iridophore lineage. To investigate *mitfa* repression, the WT simulation, which should show an initial rise followed by decline to undetectable levels of Mitfa (refer to section 4.3.1), was focused on. Four alternative GRNs were tested to identify the one that could best simulate known *mitfa* dynamics in the lineage.

4.6.1. Testing alternative models for *mitfa* repression

To minimise the number of unknown parameters and of hypotheses for which experimental testing is not currently feasible, a simple GRN was first considered, which proposed *tfec* mediated repression of *mitfa* in the iridophore lineage (Fig. 4.23 A). This relationship was modelled according to published guidelines for modelling non-competitive repression (Greenhill et al. 2011). The differential equation describing *mitfa* expression dynamics was as follows:

$$\frac{d[M^P]}{dt} = g_M \times [r_S^M] \times (1 - [r_T^M]) - d_M \times [M^P] \quad (48)$$

where

$$[r_S^M] = \frac{[S^P]}{Kd_S^M + [S^P]} \quad (49)$$

$$[r_T^M] = \frac{[T^P]}{Kd_T^M + [T^P]} \quad (50)$$

and

$$Kd_S^M = 0.2\ nM$$

$$Kd_T^M = 0.3\ nM$$

$$g_M = 0.2\ nM/h$$

$$d_M = 0.3\ 1/h$$

Parameters were chosen according to the rationale described in section 3.3.3

This version of model D (model D.1) was implemented in MATLAB to simulate gene regulatory dynamics in the WT context. It was observed that repression through *tfec* resulted in upregulation and maintenance of *mitfa* expression throughout iridophore development (Fig. 4.23 A), rather than in initial *mitfa* upregulation followed by decline to undetectable levels. Repression by *tfec* merely seemed to cause a quantitative decrease in the level reached by the *mitfa* expression plateau, although this result was disregarded, as this model is strictly qualitative.

Next, *sox10*-mediated *mitfa* repression was considered (Fig. 4.23 B). This relationship was modelled using the following equation to describe *mitfa* expression dynamics:

$$\frac{d[M^P]}{dt} = g_M \times [r_S^M] \times (1 - [r_S^M]) - d_M \times [M^P] \quad (51)$$

where

$$[r_S^M] = \frac{[S^P]}{Kd_S^M + [S^P]} \quad (52)$$

Parameters were as before for binding of *sox10* to the *mitfa* promoter. By implementing this model (model D.2), a similar result to the one from model D.1 was observed. Specifically, Mitfa concentration quantitatively decreased but qualitatively there was no change in the expression dynamics compared to model C, rendering this model biologically inaccurate (Fig. 4.23 B).

4.6.2. Monte Carlo

Before proceeding to theoretically test more complex models, confirmation was sought that the chosen parameter set was not responsible for the desired dynamics not arising in these first models. A Monte Carlo random sampling algorithm was designed, which allowed for automatic generation of random parameter combinations and screening of resulting model D outputs. The MATLAB code used to implement the following steps is shown in Appendix V.

The first step in designing the algorithm was to normalise all parameters in the model. Random values for each parameter were sampled from physiologically relevant intervals

spanning several orders of magnitude (table 4.2). For each parameter, a minimum and a maximum value were set. As shown in table 4.2, maximum expression rates (g) and protein degradation rates (d) were chosen to span 3 orders of magnitude. Dissociation constants at equilibrium (K_d) were also sampled across 3 orders of magnitude, from 0.002 nM (high binding affinity) to 2 nM (low binding affinity). Beyond 2 nM activation of gene expression was unlikely to occur according to tests performed in order to derive the set of parameter values for the model (refer to section 3.3.3.1). In every model tested so far, the initial values for Ltk, Mitfa and Pnp4a concentrations were 0 nM and were positive for Tfec (0.25 nM), Sox10 (0.4 nM) and Ltk ligand (1 nM) concentrations (refer to section 3.3.3.3). When genes were known not to be expressed, the chosen minimum and maximum values equalled 0 nM. Because *sox10* and *tfec* were shown to be expressed in the specified region of interest at 18 hpf (refer to section 3.3.3.3), Sox10 and Tfec initial values were varied between 0.01 nM and 10 nM. The Ltk ligand expression pattern has not been characterised to date, therefore initial values were sampled between 0 nM and 10 nM. k_{on} and k_{off} rates remained unchanged in these experiments (min value = max value).

Table 4.2. Parameter values used when implementing models A-D and minimum/maximum parameter values used for normalisation in the Monte Carlo algorithm.

Parameter	Values used	Min. value	Max. value
Maximum expression rate (g) (nM/h)	0.2	0.02	20
Protein degradation rate (d) (h^{-1})	0.1, 0.25, 0.5	0.05	50
Dissociation constant at equilibrium (K_d) (nM)	0.1, 0.2, 0.3	0.002	2
Initial values (nM)	0, 0.25, 0.4, 1	0, 0.01	0, 10

A total of 28 constant parameters were included in models D.1 and D.2. The parameter vector **θ** (**th**) was thus introduced to sample 28 uniformly distributed, independent random numbers from the interval (0,1) using the MATLAB command 'rand':

$$\boldsymbol{th} = rand(1,28) \tag{53}$$

Therefore, the equation used to scale the parameters in the model was

$$\boldsymbol{a} = \boldsymbol{a_{min}} + \boldsymbol{th} \times (\boldsymbol{a_{max}} - \boldsymbol{a_{min}}) \tag{54}$$

where $\boldsymbol{a}$ was a row vector containing the model's parameter set, which was determined by the random numbers drawn using the 'rand' command. The row vectors $\boldsymbol{a_{max}}$ and $\boldsymbol{a_{min}}$ contained the maximum and minimum parameter values, respectively. The system of ODEs describing model D.1 was solved for each randomly drawn parameter set $\boldsymbol{a}$. A for-loop dictated the number of times (N) that a set of 28 random parameters would be drawn. For each draw, the accuracy of the model output regarding Mitfa concentration dynamics was scored based on the difference between the maximum value reached by Mitfa along the y-axis and the minimum positive value at the plateau stage:

$$score\ (S) = \frac{[M^P]_{max}}{[M^P]_{min}} \tag{55}$$

Scores obtained from all draws were categorised in the first column of vector **S** and each was followed by the parameter values ($\boldsymbol{a}$) that were used to generate it (data not shown). The scores were sorted in descending order, with maximum scores appearing first. When N number of draws were completed, the parameter set used for the maximum score was used to plot the relevant model output.

Monte Carlo was applied for models D.1 and D.2. For each model, N = 10,000 random parameter combinations were successively implemented and outputs associated with top scores were analysed (Fig. 4.24). If parameter sets that produced biologically relevant behaviour of Mitfa existed in either of the two models, namely initial upregulation followed by subsequent decline, these sets would score highly. Results suggested that models D.1 and D.2 were both unable to predict the experimentally observed expression dynamics in the iridophore GRN, as none of the highest scoring parameter sets generated the required predictions (Fig. 4.24). In particular, an initial rise followed by downregulation of Mitfa in the iridophore lineage was only observed when iridogenic gene concentrations also dropped to undetectable levels.

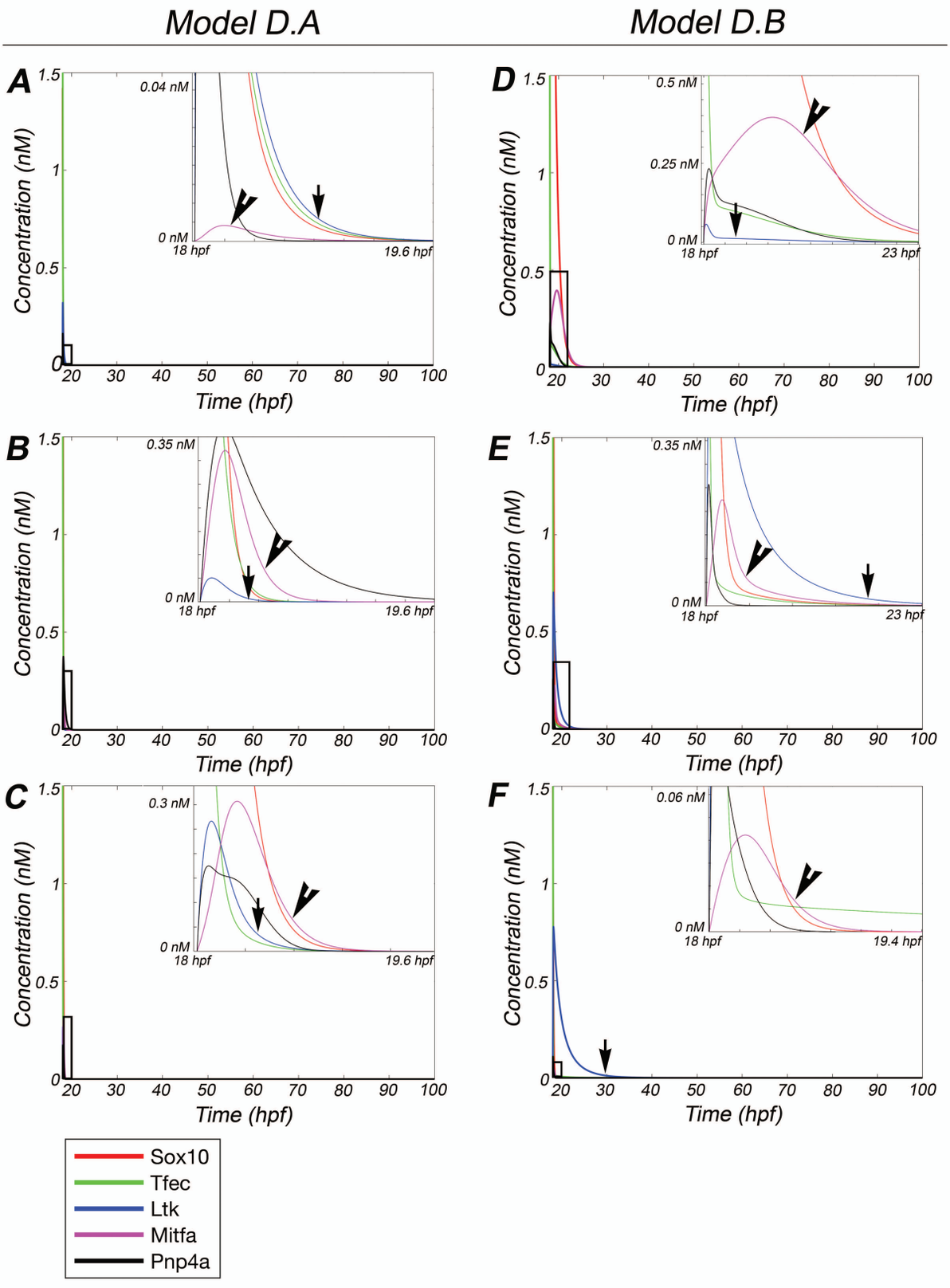


Figure 4.24. In models D.1 and D.2, Mitfa concentration displays desired behaviour only when concentrations of iridogenic genes decline over time. Output plots associated with top three Monte Carlo scores for model D.1 (A, B, C) and model D.2 (D, E, F). In each graph, rectangles correspond to positions of enlarged outputs. High-scoring Mitfa dynamics (arrowheads) require downregulation of iridogenic genes (arrows point at Ltk concentration).

4.6.3. Introducing Factor R

Since the simplistic repression models D.1 and D.2 were unable to predict experimentally observed dynamics, an unknown iridophore-specific factor R able to repress *mitfa* was introduced (Fig. 4.23 C). Because factor R functions in the iridophore lineage, it was assumed that *tfec* directly or indirectly regulated its expression (model D.3). To implement this model, the following differential equations were used to describe factor R (FR) expression dynamics (56) and *mitfa* dynamics (57):

$$\frac{d[FR^P]}{dt} = g_R \times [r_T^R] - d_R \times [FR^P] \quad (56)$$

$$\frac{d[M^P]}{dt} = g_M \times [r_S^M] \times (1 - [r_R^M]) - d_M \times [M^P] \quad (57)$$

where

$$[r_T^R] = \frac{[T^P]}{Kd_T^R + [T^P]} \quad (58)$$

$$[r_S^M] = \frac{[S^P]}{Kd_S^M + [S^P]} \quad (59)$$

$$[r_R^M] = \frac{[FR^P]}{Kd_R^M + [FR^P]} \quad (60)$$

Parameters in this model were as follows:

$$Kd_S^M = 0.2\ nM$$

$$Kd_R^M = 0.3\ nM$$

$$g_M = 0.2\ nM/h$$

$$d_M = 0.3\ h^{-1}$$

$$Kd_T^R = 0.3\ nM$$

Different values for Kd_T^R, the parameter constant describing the affinity of *tfec* for factor R, were manually tested (Fig. 4.23 C). Results indicated that, setting $Kd_T^R = 0.3\ nM$

resulted in a more noticeable difference between the maximum and the minimum Mitfa value along the y-axis ($max/min = 2.1$), compared to using lower binding affinity ($Kd_T^R = 1\,nM$), which produced a ratio of $max/min = 1.6$. Monte Carlo testing of this model is pending, but these data suggested that presence of a repressive factor R predicted experimentally established *mitfa* expression dynamics more accurately than models D.1 and D.2 did. Despite this, the required decline of Mitfa to undetectable levels was not yet apparent.

Finally, another alternative theoretical model was derived, according to which factor R was only upregulated in iridoblasts following expression of both *tfec* and *ltk* (model D.4). The two factors were set to regulate factor R via an AND gate, according to the following ODE:

$$\frac{d[FR^P]}{dt} = g_R \times [r_T^R] \times [r_L^R] - d_R \times [FR^P] \tag{61}$$

where

$$[r_T^R] = \frac{[T^P]}{Kd_T^R + [T^P]} \tag{62}$$

$$[r_L^R] = \frac{[L^P]}{Kd_L^R + [L^P]} \tag{63}$$

This model was implemented using high affinity binding constants for *tfec* and *ltk* -dependent factor R activation, as well as for factor R-dependent *mitfa* repression (in all cases $K_d = 0.1\,nM$). As shown in figure 4.23 D, this model D variant was able to predict the highest ratio of $max/min = 4.9$, while all iridogenic genes displayed biologically accurate expression dynamics (Fig. 4.23 D).

Conducting the above theoretical tests enabled derivation of testable hypotheses regarding the wiring of the iridophore GRN, so that *mitfa* upregulation would be prevented in iridophores. Identification of the proposed repressive factor R is required for experimental testing of the presented hypotheses.

Model D

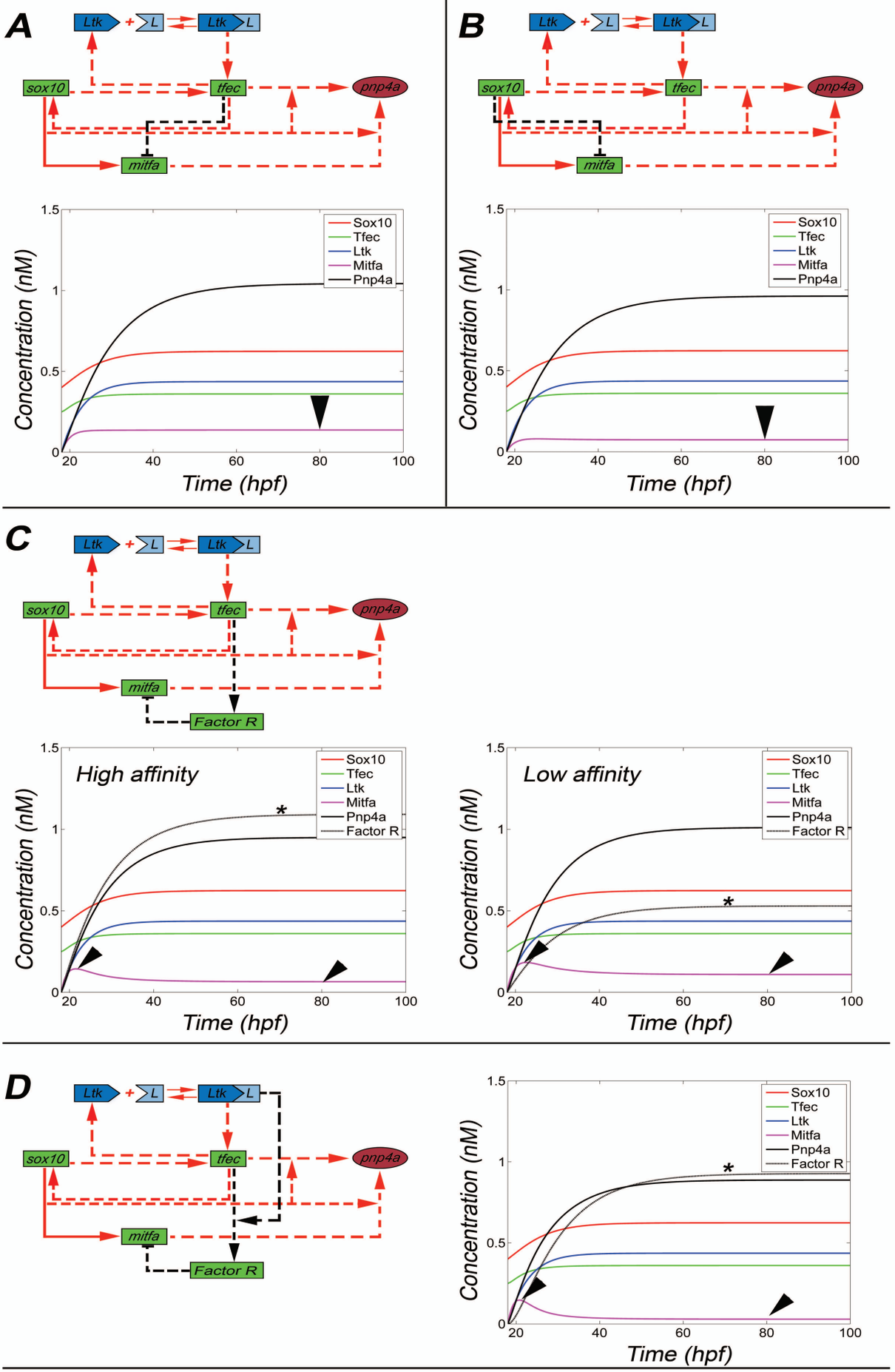


Fig. 4.23

Figure 4.23. Theoretically testing alternative models to repress *mitfa* in the iridoblast. Alternative models D.1 (A), D.2 (B), D.3 (C) and D.4 (D) were sequentially tested to achieve the most biologically accurate prediction regarding *mitfa* dynamics (arrowheads) during iridophore development, meaning an initial rise of Mitfa concentration, followed by decline to undetectable levels. Models D.1 (A) and D.2 (B) propose that *mitfa* repression is *tfec* and *sox10*-mediated, respectively. In both cases the predictions did not show the desired behaviour (arrowheads). Model D.3 (C) introduces an unknown *tfec*-dependent repressive factor R. Predictions using different dissociation constants to describe *tfec*-dependent factor R regulation (high and low affinity diagrams, asterisks) suggest that the difference between Mitfa maximum and minimum values (arrowheads) can be enhanced with the appropriate parameter sets, although the desired pattern featuring decline to undetectable levels was not obtained. Model D.4 (D) added the notion that both *tfec* activation and Ltk signalling are required for factor R upregulation. The maximum/minimum difference in Mitfa concentration was further increased (arrowheads), while expression dynamics of iridogenic genes were predicted appropriately.

4.7. Discussion

This chapter aimed to describe the iterative process of mathematical modelling and experimental testing, which led to the generation of a more sophisticated network describing the molecular basis of iridophore fate segregation from NCCs. In particular, it was suggested that *sox10* and other currently unknown factors were required for *tfec* and *mitfa* dependent *pnp4a* upregulation to take place. It was further shown that *sox10* expression was sustained in the developing iridophore lineage and that this was likely facilitated by a positive feedback loop formed between *tfec* and *sox10* in iridoblasts. Moreover, the presence of an iridophore-specific *mitfa* repressor (factor R) was proposed.

4.7.1. MiT factors in pigment cell development

The MiT factor Mitfa has been shown to be crucial for pigment cell development. In zebrafish, investigation of *mitfa* mutant phenotypes demonstrated the gene's role as the master regulator of the melanophore lineage, owing to its expression being both necessary and sufficient for melanophore development (J A Lister et al. 1999). Consistent with a role in the melanophore lineage, *mitfa* is co-expressed with melanophore differentiation genes from early stages of NC derivative specification in WT embryos (J A Lister et al. 1999). More recently, co-localisation analyses of *mitfa* with the proposed iridophore marker, *pnp4a*, indicated that, in fact, *mitfa* positive cells label a common progenitor between the melanophore and the iridophore lineage, the melanoiridoblast (Curran et al. 2010). Moreover, *mitfa* mutant lines display complete lack of melanophores, accompanied by 40% increase in iridophores (J A Lister et al. 1999), suggesting that, upon loss of *mitfa* function, the common progenitor might become biased to adopt an iridophore fate.

In contrast to *mitfa, tfec* is expressed from early stages of NCC development, when it does not solely mark iridophore precursors but expression is instead detected along the entire premigratory NC domain (Lister et al. 2011; this work). In particular, analysing *tfec* expression over a time course using conventional *in situ* hybridisation revealed that the transcript first reached detectable levels at 13 hpf in the cNC domain (James A Lister et al. 2011). At 15 hpf, the NC specification factor, *sox10*, is expressed throughout the cNC and tNC domains (Stewart et al. 2006), while at the same stage *tfec* expression was strongly detected in the cNC but also emerging in the tNC domain (James A Lister et al. 2011). By 18 hpf, both *sox10* and *tfec* label the majority, if not all, of tNCCs, but also cNCCs (Dutton et al. 2001; Lister et al. 2011; this work). These patterns suggest that *tfec* plays a role during NC development, either as part of the GRN functioning to establish the NCC population (refer to section 1.3.2), or immediately downstream of the network. Nevertheless, little is currently known with regard to potential *tfec* functions

during NC specification and the gene has not been considered a member of the NC specification GRN (M. Simões-Costa et al. 2015). Interestingly, over the course of this project, *tfec* expression at 18 hpf has been investigated upon loss of function of genes with recognised roles during NC establishment. In particular, *sox10*, *foxd3* but also *sox9b* and *tfap2a* mutants were examined and, strikingly, *tfec* expression within premigratory NCCs remained unaffected (data not shown). This observation reinforced the notion that *tfec* is part of the NC specification GRN and under the control of redundant activatory effects, similarly to what has been proposed for other transcription factors in the network (M. Simões-Costa et al. 2015). Following 18 hpf, *tfec* expression becomes progressively downregulated in a rostro-caudal manner, and has been shown to only be maintained in presumptive iridoblasts (Lister et al. 2011; this work). However, the mechanism by which *tfec* is regulated during this transition remain elusive.

4.7.1.1. Cross-repression between *mitfa* and *tfec* during specification of pigment cell fates

Results described in this chapter provide evidence of a repressive effect of *mitfa* on *tfec* during early stages of NC specification. Specifically, it was shown that in *mitfa* mutant embryos, contrary to the WTs, *tfec* fails to be repressed in NC-derived cells along the premigratory domain of the dorsal trunk and in both medially and laterally migrating cells. One explanation for this phenotype is that *mitfa*-dependent transcriptional regulation normally represses *tfec* expression, likely in an indirect manner. This hypothesis is supported by the expression patterns of the two genes. *tfec* is initially expressed in the premigratory NC domain at stages when *mitfa* expression is not detectable. As *mitfa* expression becomes detectable along the dorsal trunk in a rostro-caudal manner, *tfec* is progressively restricted towards the posterior of the embryo. Moreover, co-expression studies presented here demonstrated that in cells where the two genes are co-expressed during early stages of specification, one consistently appears to be more abundant than the other. Collectively, these data highlight a likely key molecular event occurring during melanoiridoblast fate choice. Results support a repressive effect of *mitfa* on *tfec* expression and, potentially, a similar effect of *tfec* on *mitfa* expression in the iridophore lineage. However, the latter has yet to be confirmed and requires further investigation using the newly established *tfec* mutant lines. Until further evidence becomes available, the effect of another, currently unknown factor inhibiting *mitfa* in the melanoiridoblast and allowing maintenance of *tfec* expression cannot be dismissed.

4.7.1.2. MiT factors potentially facilitate chromatophore specification

It has been previously shown that Tfec is able to heterodimerise with other MiT family proteins, including Mitf and potentially Tfeb and Tfe3, to exert its transcriptional effects (Hemesath et al. 1994; Pogenberg et al. 2012). Interestingly, different ratios of MiT

factors have been suggested to activate diverse transcriptional programs (Kuiper et al. 2004). The notion that Mitfa/Tfec heterodimers can potentially exert transcriptional effects different than those triggered by Mitfa/Mitfa or Tfec/Tfec homodimers is very interesting in the context of dynamic fate restriction processes that progressively guide NCCs to adopt distinct fates.

More specifically, as noted in the previous section, the expression patterns of the two genes do not overlap extensively between 18-24 hpf, when fate restriction processes are thought to primarily take place. Since turnover of protein occurs less rapidly than of mRNA, it is likely that Mitfa and Tfec TFs co-exist in NCCs undergoing specification for a limited period of time, during which heterodimers could be formed. Later in development, exclusion of Tfec from the melanophore lineage and of Mitfa from the iridophore lineage might indicate that homodimers might play specific roles in fate-restricted, possibly unipotent, progenitors. In order to elucidate such effects, possible experiments might include single cell transcriptomics, which would allow for identifying progenitors co-expressing *mitfa* and *tfec* and for relating this feature to the expression profile of the corresponding cells. These profiles could then be compared to those of cells only expressing one or the other TF. It should also prove valuable to assess *in vivo* binding targets of Mitfa or of Tfec in different contexts, by isolating populations of cells of interest using transgenic lines and performing ChiP-Seq at different stages of development.

4.7.1.3. Cells other than mature iridophores express *tfec* at differentiation stages

Previous studies first demonstrated that *tfec* is a marker of mature iridophores during differentiation stages (following 42 hpf) (James A Lister et al. 2011). Results presented in this chapter raised questions regarding the nature of the cells comprising the differentiated iridophore population. Persistence of *tfec* positive cells forming a pattern reminiscent of mature iridophores along the dorsal stripe of presumptive *tfec* mutants suggested the existence of mature iridophore subpopulations, in which different GRNs may function. However, this is improbable taking into account that escaper iridophores can be very rarely observed in *tfec* mutants. It is more likely that the cells in question correspond to NC-derived progenitors, for example chromatoblasts, which co-localise with mature iridophores but are not yet specified since they do not express iridophore lineage markers, such as *ltk*. These results await confirmation with larger sample sizes. Whether these cells are NC-derived could potentially be determined by generating *sox10:*Cre*;tfec*$^{-/-}$ transgenic lines. The identity of the *tfec* positive cells could be further characterised by co-labelling mutant embryos for *tfec* expression and markers of other cell types, such as *foxd3* for NCC progenitors or *mitfa* for the melanoiridoblast using

RNAscope. In addition, RNAscope could be used to determine whether *tfec* positive cells are positive for phosphohistone H3, thus indicating active proliferation.

Overall, *tfec* appears to function in several stages of NC and pigment cell development, while some of its effects are likely exerted in cooperation with, or through repressive feedback with the melanophore master regulator, *mitfa*. With regards to the iridophore GRN, the effect of loss of *tfec* function on *sox10* expression throughout iridophore development remains to be addressed. Expression studies on more time points (30 hpf and 36 hpf) are required to fully characterise the effects on *mitfa* positive melanoblasts and on NC derivatives expressing *pnp4a*, while co-expression studies will determine whether *mitfa* can be successfully repressed in iridophore precursors upon loss of *tfec* function. Interestingly, live imaging of melanophores, as presented in this chapter, highlighted delay of melanogenesis in *tfec* mutants and a potential increase in differentiated melanophores. Future studies should thoroughly characterise these effects via scoring numbers of melanophores in WT and *tfec* mutant embryos.

4.7.2. Activation of *pnp4a* in the iridophore lineage

Previous studies described *pnp4a* as an early marker, specific to the iridophore lineage (Curran et al. 2010). Surprisingly, in chapter 3 not only *tfec*, but also *mitfa* were identified as either direct or indirect upstream regulators of *pnp4a*. In the process of evaluating predictions from model A, it was demonstrated that upon loss of *sox10* function, *tfec* alone, which was still expressed in trapped premigratory precursors, was unable to upregulate *ltk*. Similarly, gradual elimination of the activatory effect of *mitfa* on *pnp4a* expression was linked to concomitant downregulation of *sox10* as melanophore development progresses (Greenhill et al. 2011). Therefore, MiT factor-dependent *pnp4a* regulation was postulated to only occur in the presence of functional *sox10*. In order to test this hypothesis, co-overexpression experiments were attempted, during which *sox10*/*mitfa* and *sox10*/*tfec* transcripts were injected simultaneously in single-cell stage embryos.

Co-overexpression studies were often difficult to interpret as a result of large error bars and effects in gene expression levels observed in negative controls when compared to reference samples. These effects might result from small variations in the developmental stage of embryos used for these experiments, which could be caused by frequently observable developmental delays, or even developmental arrest, in injected samples. In addition, variable emission of fluorescence due to increased primer dimer formation during q-RT PCR amplification, owing to low or absent expression of target genes in embryos at this stage, could be responsible for slight apparent increases in expression levels (for example 2-fold). Therefore, small increases in gene expression levels were cautiously interpreted as they could correspond to experimental artefacts.

Strikingly, in all overexpression studies, *pnp4a* upregulation failed, while the gene is expressed in the iridophore lineage (Curran et al. 2010; this work) in a *tfec*-dependent manner, according to loss of function studies (this work). Furthermore, experiments presented here suggested that *mitfa* also activated *pnp4a* transcription during early stages of melanophore specification. Thus, the inability of both *tfec* and *mitfa* to ectopically activate *pnp4a* expression, in the presence or absence of *sox10*, was unexpected. This result could have arisen because the effects of overexpression were assayed in a non-physiological context. Firstly, essential cofactors or signals might be missing from 6 hpf and 10 hpf embryos, although they are present in NCCs later in development. Secondly, in early embryos, unknown effects could prevent the overexpressed factors from activating *pnp4a* expression. Specifically, since endogenous *pnp4a* was found by q-RT PCR to be actively expressed between 6-10 hpf in negative controls (this work, data not shown), it is plausible that additional activation through ectopically expressed TFs is prevented, or obscured. To overcome the context issue, overexpression of cDNA, followed by assessment of *pnp4a* expression levels in NCCs at approximately 18-24 hpf should be performed. Alternatively, stable overexpression of appropriate genes, for example through insertion of CMV-driven sequences into the genome, is an option. To address the second hypothesis, overexpression of TFs of interest co-injected with a *pnp4a*:GFP fusion gene construct could indicate whether the chromatin structure or the presence of endogenous *pnp4a* activators or repressors prevent binding of exogenously supplied activators.

Another point to consider is whether *pnp4a* activation can only be achieved by a *tfec* isoform different than the one injected. Three different *tfec* isoforms, encoded by alternative promoters, bearing distinct first exons and functioning in different tissues have been identified in zebrafish (Mahony et al. 2016). Moreover, heterodimerisation between Tfec and other MiT family proteins is important for the TF to exert its transcriptional effects (Hemesath et al. 1994; Pogenberg et al. 2012). To further complicate matters, combinations of specific ratios of different MiT factors have been suggested to activate diverse transcriptional programmes, potentially by altering heterodimerisation tendencies (Kuiper et al. 2004). Taking these mechanistic considerations into account, achieving upregulation of *pnp4a* might require overexpression of different *tfec* isoforms or of combinations of MiT factors, most likely *mitfa/tfec*, since both genes are expressed in the NC and mutants for either gene lack *pnp4a* expression (this work). In the case of future experimental confirmation of the cooperativity between *mitfa/tfec* for *pnp4a* activation, rather than between *tfec/sox10* or *mitfa/sox10,* the role of *sox10* as essential regulator of *pnp4a* activation is not abrogated, since it is responsible for upregulation of both *tfec* and *mitfa* in the GRN and, therefore, remains indirectly responsible for *mitfa/tfec* cooperation.

Interestingly, in all experiments, overexpression of *tfec* resulted in statistically significant repression of *pnp4a* at 10 hpf. Additional experiments are required to confirm whether this result is biologically meaningful, or if it is simply an artefact associated with overexpression. According to loss of function studies, *tfec* function is important for *pnp4a* upregulation in NCCs undergoing specification, which argues against *tfec*-mediated inhibition of *pnp4a* expression. It might, however, be the case that this effect is context dependent. For instance, *tfec* might interfere with transcriptional machinery responsible for activating *pnp4a* in the early embryo, thus resulting in repression of *pnp4a*. Another explanation is that Tfec homodimers, which most likely occur as a result of overexpressing *tfec* in the early embryo, may act as repressors of *pnp4a*. It is worth testing whether Tfec heterodimers, either with Mitfa or other bHLH TFs, exert an activatory effect on *pnp4a* regulation.

4.7.3. Loss of *ltk* dependency in model C

When attempting to overcome the loss of *ltk* dependency in model C, different alternative GRNs were tested theoretically, and underlying hypotheses were evaluated experimentally using overexpression studies. All three proposed alternatives, C.1, C.2 and C.3, were discarded on the basis of theoretical testing not being in agreement with the presented overexpression data. Specifically, predictions from mathematically modelling the three networks suggested that C.3 outputs were consistent with observations from loss of function studies. However, evidence from overexpression experiments argued against the validity of the genetic interactions depicted in C.3 and suggested that model C.2 was more biologically relevant.

Briefly, overexpression studies suggested that (i) *tfec*-mediated activation of *sox10* likely occurred independently of Ltk signalling, (ii) *ltk* overexpression was not sufficient to activate *tfec* expression, and (iii) *sox10/ltk* co-overexpression did not observably upregulate *tfec* expression. When considering these results, potential false positives or negatives, which may have resulted from variation of gene expression levels between negative controls, should not be disregarded. Thus, additional experiments will be useful to provide further support for the presented data.

Regarding observation (i), it should be considered that alternative RTK signalling present during early embryogenesis might be responsible for mediating *tfec*-dependent *sox10* activation, thus obscuring *ltk* dependency in this interaction. Moreover, *tfec*-dependent endogenous *ltk* activation could be responsible for activating *sox10*, even when *npm-ltk* mRNA was not co-injected.

To eliminate these options and definitively test whether *sox10* activation requires both *tfec* and *ltk*, a generic RTK inhibitor could be used to eliminate any transcriptional effects

occurring through RTK signalling. To that end, a useful candidate for co-overexpression is the E3 ubiquitin protein ligase, c-Cbl. c-Cbl has been identified as a dominant inhibitor of RTKs and of proteins interacting with RTKs, by tagging them for degradation (Thien and Langdon 2005). In zebrafish, loss of *c-cbl* function has been associated with phenotypes analogous to human disease, mediated by signalling from the RTK *fms*-like tyrosine kinase 3 (Flt3) (Peng et al. 2015). Furthermore, the compound sunitinib has been reported to successfully inhibit effects of a number of RTKs, including Kit (Roskoski 2007). A variety of additional RTK inhibitors have been identified, most of which function to prevent signalling through specific RTKs (Ledda and Paratcha 2007).

Observations (ii) and (iii) could be due to overexpressing *ltk* in a non-physiological context, where important cofactors and cues are missing. Transient overexpression of cDNA encoding factors of interest, or stable overexpression by inducing integration of a construct within the genome, followed by assessment of transcriptional outcomes during NC formation would be useful to rule out context-specific effects.

4.7.4. Monte Carlo

Over the course of theoretical testing and experimentation, a Monte Carlo random sampling algorithm was generated. This allowed for more rigorous exploration of the parameter space in order to establish whether simple models were able to produce desired predictions, consistent with experimental observations, prior to generating more complex and difficult to test models with additional assumptions. These preliminary attempts used a simple scoring method to highlight appropriate models, by which desirable, high-scoring models were the ones with a large difference between the maximum and the subsequent minimum value reached by Mitfa, meaning that *mitfa* first became upregulated and subsequently declined to undetectable levels.

Although this scoring method addressed certain needs in this study, it will require improvements. Specifically, scoring should take into account the requirement that iridophore-specific genes must remain upregulated throughout iridophore development. Therefore, parameter sets that result in downregulation of all genes in the network would be penalised. Importantly, the current method of scoring can only be applied if Mitfa follows a simple pattern of peaking once and subsequently declining to a plateau stage. However, a more advanced model or specific parameter sets could result in several peaks or even oscillatory behaviour of Mitfa, in which case local minima would interfere with the scoring mechanism. Hence, it is important for the scoring equation to include a penalising term, which would result in successive local maxima and minima scoring low.

4.7.5. Summary

In this chapter, both mathematical modelling and experimentation were used to progressively refine and expand the experimentally described model A to generate a more sophisticated model D describing the molecular basis of iridophore fate segregation from NCCs. Model D included unknown factors and proposed previously unacknowledged, experimentally testable interactions, some of which require further evaluation. However, some of the predictions from model D remain inconsistent with known biology. Therefore this model can be developed in the future to further improve the iridophore GRN.

Factor R arose as the repressor of the melanophore master regulator, *mitfa*, in the iridophore lineage and was found to be an essential component of the mathematical model. In the next chapter, experimental characterisation of factor R was attempted. Identification of this predicted player of the network will then allow for thorough evaluation of the GRN and likely shed light on the melanoiridoblast fate decision process.

5

Investigation of candidates for factor R

5.1. Introduction

Chapter 4 outlined the iterative process of improving the iridophore GRN using mathematical modelling and experimental genetics. This modelling approach predicted that a repressor, termed factor R, was required to suppress *mitfa* in the iridophore lineage, following its direct activation by *sox10*. This chapter focuses on the experimental approach taken towards the identification of factor R. Specifically, previously conducted studies indicated that the transcriptional repressors *foxd3* and *id2a* are promising candidates for the role of factor R. Loss of function studies, expression and co-expression assays were employed in order to elucidate the potential role of the two genes in the iridophore GRN.

The first candidate tested for the role of the *mitfa* repressor was *foxd3*. The gene has been previously suggested to repress *mitfa* in several contexts, including *in vitro* models (Thomas and Erickson 2009) as well as *in vivo* in zebrafish embryos (Ignatius et al. 2008; Curran, Raible, and Lister 2009; Curran et al. 2010). Studies using zebrafish models indicated that *foxd3* was important in the melanoiridoblast fate switch process, during which it was reported to repress *mitfa* allowing iridophores to become specified (Curran et al. 2010). However, requirement for *foxd3* has been described in different contexts. Mammalian epiblast cells and ESCs reportedly require Foxd3 for their maintenance and establishment, respectively (Hanna et al. 2002). Moreover, this TF plays an important role in determining and maintaining the NC fate in *Xenopus* embryos (Sasai, Mizuseki, and Sasai 2001; Pohl and Knöchel 2001). It is therefore postulated that *foxd3* functions in multipotent NC progenitors to repress differentiation genes, such as *mitfa*, in the presence of appropriate cofactors. Specifically, avian PAX3 has been shown to be required for FOXD3 to repress MITF expression (Thomas and Erickson 2009).

Based on the hypothesis that *foxd3* represses *mitfa* in NC derivatives, it is predicted that loss of *foxd3* function would result in reduction of iridophores accompanied by an

increase in *mitfa* positive developing and mature melanophores. However, in zebrafish embryos morpholino-mediated elimination of *foxd3* function caused a reduction in iridophores but did not lead to an increase in cells expressing melanophore markers downstream of *mitfa* (James A Lister et al. 2006). Furthermore, detailed analysis of the effects of the *mos*^{m188} as well as the *sym1 foxd3* mutant alleles revealed a delay in melanogenesis (Montero-Balaguer et al. 2006; Stewart et al. 2006). Differentiated melanophore numbers were found to be mildly decreased in the former line (Montero-Balaguer et al. 2006) and unchanged in the latter (Stewart et al. 2006). In both *mos*^{m188} and *sym1* mutant embryos iridophores were reportedly decreased. These data argued against a fate switch of iridophores towards the melanophore lineage that would be expected from the melanoiridoblast model (Curran et al. 2010). Conversely, a novel *foxd3* mutant line, gt(*foxd3*-cherry)^{c110aR}, which resulted from the insertion of a transgene within the *foxd3* coding region, revealed a striking increase in melanophore numbers, in addition to a reduction in mature iridophores (Hochgreb-Hägele and Bronner 2013).

A step towards understanding the basis of these conflicting results with regards to melanogenesis phenotypes in *foxd3* mutant embryos was to consider the nature of the respective mutant lines. The mutation in the *mos*^{m188} line has yet to be identified. Thorough investigation of the *foxd3* coding region did not reveal any lesions that would account for the severe craniofacial, neural and pigment cell defects (Montero-Balaguer et al. 2006). It was, thus, concluded that the mutation likely lies within a regulatory region controlling *foxd3* expression, although whether such a mutation also affects other genes remains undetermined.

The *sym1* mutation was identified as a nonsense mutation, resulting in truncation of the *foxd3* transcript at the 186th amino acid residue, out of 371 residues in total (Stewart et al. 2006). This allele was claimed to be a null, on the basis of its recessive nature and the loss of crucial TF domains, including the nuclear localisation signal. However, only one out of the three helices which are important for *foxd3*-mediated transcriptional regulation is disrupted, while the N-terminal acidic domain of the TF remains intact (Stewart et al. 2006). Therefore, it might be the case that *foxd3* still exerts a regulatory effect on transcription.

Finally, the gt(*foxd3*-cherry)^{c110aR} transgenic line results in disruption of the *foxd3* transcript at the 100th amino acid, thus eliminating all the major functional domains of the protein. It remains to be shown whether the increase in melanophore number observed in the transgenic line was, indeed, the result of complete abolishment of *foxd3* function, or a non-specific effect of the inserted construct.

Here, the newly generated mutant allele *foxd3*^{sa20726} was used to characterise the effects of loss of *foxd3* function in the development of melanophores and iridophores. The single

nucleotide change in these mutants resulted in premature truncation of the protein at the 40th amino acid residue, resulting in loss of the major conserved regions of the TF, which are located after the 90th amino acid. This mutant proved to be a valuable tool to complement previous studies on the effects of *foxd3* during chromatoblast specification and differentiation. As a result of loss of function and co-expression studies, *foxd3* was rejected for the role of factor R.

Following exclusion of *foxd3*, the potential role of *id2a* (inhibitor of differentiation 2a) as the repressor of *mitfa* in iridophores is investigated. Id family proteins contain an HLH motif, but not a basic DNA-binding domain and have been known to generate heterodimers with E-box binding bHLH TFs, thus preventing them from binding to DNA (Ling, Kang, and Sun 2014). *mitfa* is a bHLH E-box factor, although to our knowledge its repression by Id family TFs has not been reported to date. Intriguingly, morpholino-mediated knockdown of *id2a* resulted in complete lack of iridophores (F. R. Rodrigues, PhD thesis). Moreover, overexpression studies by Dr Tatiana Subkhankulova (Kelsh group) suggested that overexpression of *id2a* mRNA in single-cell stage embryos led to ectopic overexpression of *tfec* at 6 hpf. Since *id2a* functions by repressing TF activity, and considering data from chapter 4, which indicated likely repression of *tfec* by *mitfa*, these results may suggest direct or indirect regulation of *mitfa* function by *id2a* in the iridophore lineage.

To evaluate *id2a* for the role of factor R in the iridophore GRN, *id2a* expression in iridoblasts was investigated using co-expression studies. Furthermore, *id2a* mutant lines were generated using CRISPR/Cas9 technology to directly test the gene's role in repressing *mitfa* function in iridophores.

5.2. Effect of *foxd3* loss of function on the iridophore lineage

5.2.1. Genotyping *foxd3^{sa20762}*

The nature of the SNP in *foxd3^{sa20726}* mutants (C to A, pyrimidine to purine) resulted in a relatively pronounced change in the melting temperature of the surrounding genomic region. Therefore, genotyping of live *foxd3* mutant embryos was easily conducted using HRMA. A pair of primers flanking the SNP was designed (refer to Appendix II), which produced an amplicon of 142 bp (60.5% GC content). For this assay different parameter sets and alternative conditions were tested before choosing the optimal parameters (data not shown, refer to Appendix II for details on optimised experimental parameters). Using HRMA, WT, heterozygous and homozygous mutant sequences could be distinguished (Appendix Fig. III.1 A). HRMA-based genotyping was attempted for genotyping single embryos which had been subjected to *in situ* hybridisation, but the high level of chromatin

crosslinking in these samples resulted in low amplification efficiency during q-RT PCR, which led to obtaining unreliable predictions by HRMA (Appendix Fig. III.1 B).

To genotype fixed embryos, conventional PCR and restriction digest-based genotyping was attempted. Initially, the restriction enzyme *DdeI* was used to digest PCR products generated using the q-RT PCR primer set. The enzyme recognised the mutant allele, but not the WT (refer to Appendix III), and almost completely digested PCR products derived from homozygous mutant embryos, while generating faint bands in heterozygous mutants and not at all digesting WT alleles (Appendix Fig. III.1 C). To efficiently amplify this GC rich region of genomic DNA derived from *in situ* hybridisation samples, the highly sensitive KAPA Robust DNA Polymerase was used, rather than other versions of Taq polymerases (Appendix Fig. III.1 E). Attempts to de-crosslink chromatin from extracted samples in order to achieve higher amplification yield were not successful (Appendix Fig. III.1 E).

To allow for easier visualisation of genotyping results, primer sets generating larger amplicons were designed and tested (refer to Appendix II; Appendix Fig. III.1 D). The preferred primer set generated a 227 bp amplicon (61.2% GC content). Amplification of the desired GC-rich genomic region from cross-linked genomic DNA templates was still suboptimal. Further tests indicated that addition of 5% DMSO (final concentration) and 30 µg BSA for each 10 µl PCR reaction resulted in higher amplification yield (Appendix Fig. III.1 G,H), although several samples were still unable to produce amplicons. Testing different extraction and chromatin de-crosslinking protocols (data not shown) suggested that instead of the typically used KAPA Express Extract kit, a proteinase-K extraction protocol (refer to section 2.3.1) was best used (Appendix Fig. III.1 H).

5.2.2. The iridophore lineage is affected in *foxd3*^{sa20726} mutants

First, the effect of loss of *foxd3* function on the iridophore lineage was characterised using the *foxd3*^{sa20726} line. Iridophores of the dorsal and ventral stripe of 62 embryos at 4 dpf were scored. The embryos were derived from three independent incrosses between identified mutation carriers. Individual embryos were then genotyped by HRMA (Fig. 5.1 B) and the number of iridophores along the dorsal and the ventral stripe of WT, heterozygous carriers and homozygous mutants was plotted (Fig. 5.1 A). Paired t-tests were performed to establish statistical significance between mean values of each group.

The number of iridophores was not statistically different between WT and heterozygotes along the dorsal stripe (12% decrease in iridophores in the latter group), or along the ventral stripe (16.1% decrease). However, the degree of heterogeneity in the number of iridophores appeared to increase in heterozygous embryos compared to WTs. In mutant embryos, compared both to WT and to heterozygous siblings, the number of iridophores

decreased significantly. Specifically, iridophores along the dorsal stripe of homozygous mutants were reduced by 70.5% compared to WT siblings and by 66.4% decrease compared to heterozygous siblings. No statistically significant differences were observed in the numbers of iridophores along the ventral stripe between WT or heterozygous embryos and mutant siblings (Fig. 5.1 A). Specifically, there was a 10.5% decrease between iridophores of WT and mutant siblings. Again, increased variance was observed between WT embryos and *foxd3*^{sa20726} allele carrier siblings.

It was concluded that, although the previously described jaw malformation phenotype (Montero-Balaguer et al. 2006) was not observed in the majority of the *foxd3*^{sa20726} mutant embryos (data not shown), the iridophore phenotype was as previously reported in other loss of function studies (James A Lister et al. 2006; Montero-Balaguer et al. 2006; Stewart et al. 2006).

To support this finding, *in situ* hybridisation studies were conducted to detect expression of iridophore markers in embryos fixed at 36 hpf, derived from incrosses between *foxd3*^{sa20726} heterozygous carriers. At this stage, specified iridoblasts occupy the dorsal trunk and are still detectable in migratory pathways. The expression patterns of *tfec* (Fig. 5.1 C-F), and *pnp4a* (Fig. 5.1 G-J) were investigated. Scoring *tfec* and *pnp4a* positive cells along the dorsal trunk revealed a continuum of cell numbers (data not shown), rather than two discrete subsets of embryos, specifically 75% (WT and heterozygous mutants) with high and 25% (corresponding to homozygous mutants) with low numbers of iridoblasts. This result was linked to the observed high degree of variability in the numbers of live iridophores present in heterozygous embryos (Fig. 5.1 A). To interpret the data for these experiments, it was assumed that embryos with 5 or less iridoblasts along the dorsal trunk corresponded to homozygous mutants. The null hypothesis that these individuals comprised 25% of the total assessed embryos was formulated and χ^2 tests were performed. However, it remained likely that some of the presumed mutant embryos corresponded to heterozygous individuals or that a small number of homozygous mutant embryos with more than 5 iridoblasts along the dorsal trunk were mistakenly identified as WT or heterozygous.

At 36 hpf, over 3 independent experimental repeats, a total of 19/79 embryos (mutants according to Mendelian ratios; $0.8 < \text{p-value} < 0.9$) showed 5 or less *tfec* positive cells along the dorsal trunk domain. These embryos presented with an average of 3 cells in this domain. The developing LPs and the ICM maintained *tfec* expression, thus suggesting that the observed reduction was not an experimental artefact (Fig. 5.1 D,F). On average, 12 stained cells were present along the dorsal trunk of remaining embryos (Fig. 5.1 C,E). A 70.9% decrease in the number of *tfec* positive cells along the dorsal trunk of presumptive mutants was thus observable, reflecting the decrease observed in

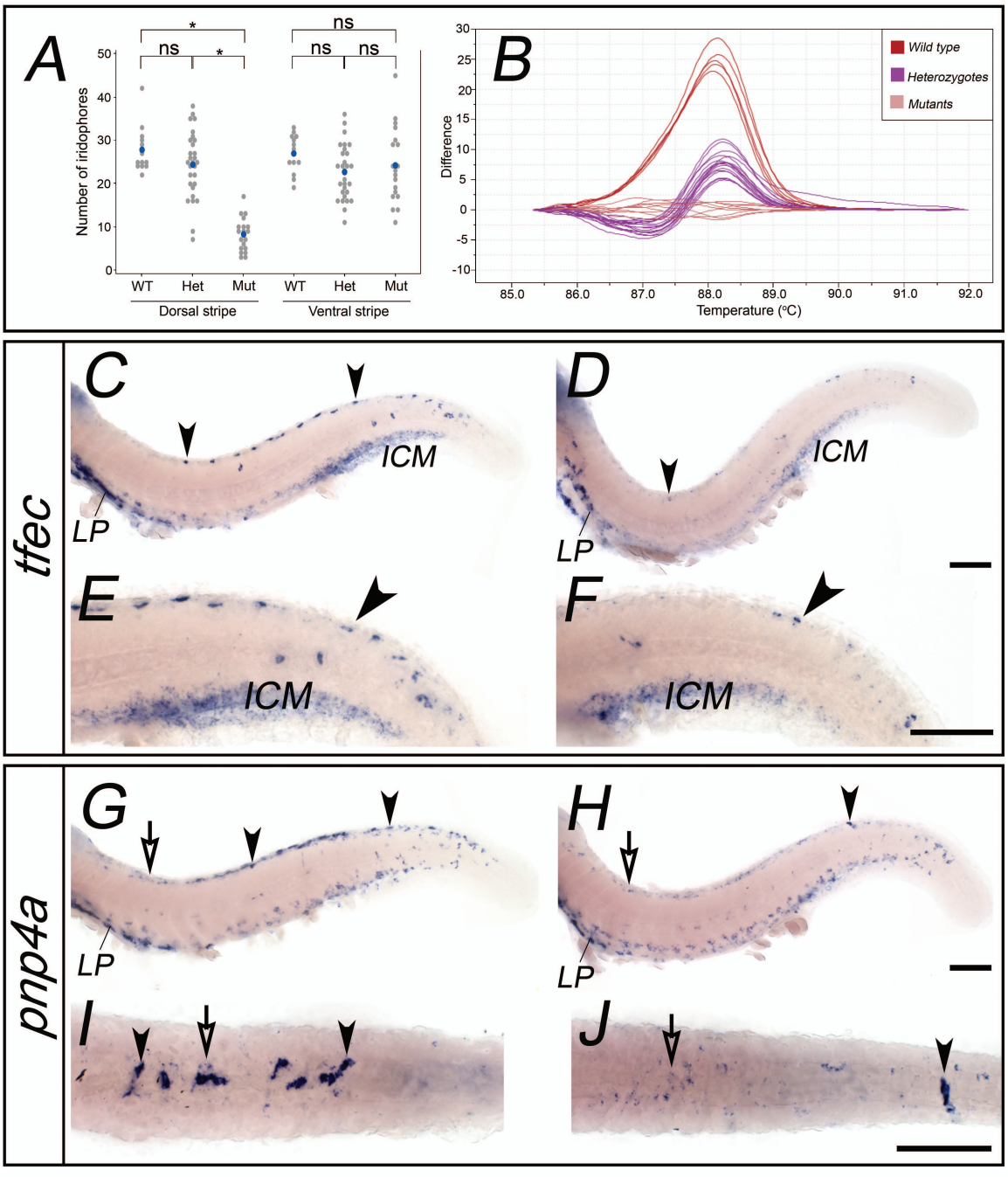


Figure. 5.1. Specified iridoblasts and mature iridophores are reduced in *foxd3*^{sa20726} mutants. (A) Numbers of dorsal and ventral stripe iridophores in genotyped WT, f*oxd3*^{sa20726} heterozygous carriers and homozygous mutant siblings. (B) HRMA-derived difference plot, used to indicate the genotypes of scored embryos. (C, E) *tfec* is expressed in specified iridoblasts along the dorsal trunk (arrowheads), in the LPs and in the ICM region. (D, F) In presumptive *foxd3* mutants, *tfec* is expressed in a reduced number of cells along the dorsal trunk (arrowheads), but expression is maintained in the LPs and appears normal in the non-NC derived ICM region. (G, I) *pnp4a* is strongly expressed in specified iridoblasts (arrowheads), weakly in presumed cells of the melanophore lineage (empty arrows) and in the LPs. (H, J) In presumptive *foxd3* mutants, *pnp4a* is expressed in very few iridoblasts (arrowheads) and appears reduced in LPs but unaffected in weakly stained presumed melanophores (empty arrows). * p-value < 0.001; ns, not significant; LP, lateral patches; ICM, intermediate cell mass. C-H: lateral views. I, J: dorsal views. Oriented with the head to the left. Scale bars: 100 µm.

live iridophores upon loss of *foxd3* function. Migrating cells were still detectable in affected individuals, suggesting that at least a subset of iridoblasts was successfully specified.

Using a probe against the *pnp4a* transcript, in a single experimental repeat 12/25 examined embryos (p-value < 0.001) presented with 5 or less *pnp4a* positive iridoblasts (2 cells on average; Fig. 5.1 H,J), while up to 19 iridoblasts were present along the dorsal trunk in the rest of the examined embryos (10 cells on average; Fig. 5.1 G,I). Thus, the dorsal trunk iridoblasts presented with a 77% reduction between presumptive WT or heterozygous and mutant embryos. The low p-value indicated that expected Mendelian ratios were not reflected in the observed number of affected embryos. This was likely due to the low sample number combined with the high phenotypic variability observed amongst heterozygous embryos, which could lead to falsely scoring a subset as potential mutants. Similarly to when *pnp4a* expression was assessed on *ltk*^{ty82} mutants (section 4.3.3), iridoblasts were identified as the prominently stained cells. Cells with weaker staining, presumably corresponding to cells of the melanophore lineage, appeared grossly unaffected by loss of *foxd3* function (Fig. 5.1 G-J). This experiment will need to be repeated to confirm a statistically significant decline of *pnp4a* positive cells in *foxd3* homozygous mutants.

Overall, these data confirmed that *foxd3* plays a role in the development of the iridophore lineage. Although the persistence of *tfec* positive migrating iridoblasts in *foxd3* mutants indicated specification took place, it remained unclear whether the observed effect was the result of partial specification defects or of a later role of *foxd3* in specified and differentiated iridophores. The latter would support the notion that *foxd3* represses expression of melanophore markers in iridophores.

5.2.3. *foxd3* mediates pigment cell specification

The role of *foxd3* during iridoblast specification was further investigated. To that effect, expression dynamics of the iridophore lineage marker *tfec* were investigated by *in situ* hybridisation at 18 hpf, 24 hpf and 30 hpf. For all experiments in this section, embryos were obtained from incrosses between identified *foxd3*^{sa20726} carriers. Therefore 25% of assessed individuals were expected to be homozygous mutants, according to Mendelian ratios.

tfec expression at 18 hpf was unaffected in N = 31 embryos analysed. The staining intensity at the premigratory tNC domain as well as the domain's degree of expansion along the rostro-caudal axis were examined (data not shown). A subset of these embryos was genotyped, which confirmed that any mild variations were due to experimental

artefacts and did not correlate with homozygosis of the mutant allele (data not shown). Therefore, no defects in the establishment of *tfec* positive NCCs were identified.

At 24 hpf, in 9/57 assessed embryos (mutants according to Mendelian ratios; 2 independent experimental repeats; 0.1 < p-value < 0.2) the premigratory NC domain expressing *tfec* was expanded from the posterior, where it was confined in WT embryos (Fig. 5.2 C), towards the anterior (Fig. 5.2 D). This pattern of *tfec* expression was similar to the one observed in *sox10*^{t3} mutants (section 3.2.2), although less pronounced. In *sox10* mutants the tNC premigratory domain was prominently stained from the hindbrain-trunk boundary to the tail, and *tfec* positive migrating cells were absent (Fig. 3.3 A,B). In *foxd3* mutants the expansion of the tNC domain was not as evident and a small number of migrating cells was still observed (Fig. 5.2 D). Presumptive *foxd3* mutants also showed an observable increase in the intensity of the *tfec* positive cNCC domain, and fewer cells were observed in the developing LPs (Fig. 5.2 C,D). To confirm that this phenotype manifested in homozygous *foxd3*^{sa20726} mutants, but not in WT siblings, 35 random embryos of the 57 assessed in total were imaged and then genotyped. As a result, it was confirmed that only homozygous mutants displayed the observed phenotype (data not shown).

At 30 hpf, 5 or less *tfec* positive cells corresponding to specified iridoblasts were observed along the dorsal trunk of 17/52 embryos (mutants according to Mendelian ratios; 2 independent experimental repeats; 0.2 < p-value < 0.3; Fig. 5.2 G,H). Between presumptive WT or heterozygous embryos and mutant siblings, there was a 71.9% decrease in cells along the dorsal trunk, but only a 1.8% decrease in migrating cells. In all embryos, *tfec* expressing premigratory NCCs of the posterior tail were present, although perhaps decreased in presumptive mutants (Fig. 5.2 G,H). *tfec* expression was reduced in developing LPs in these embryos. Similar observations were made when *ltk* expression was assessed in *foxd3* mutants (data not shown). Of the examined embryos, 10/26 showed 5 or less *ltk* positive specified iridoblasts along the dorsal trunk (mutants according to Mendelian ratios; single experimental repeat; 0.1 < p-value < 0.2). The average number of cells in putative mutants was reduced by 66.7% compared to the average number in presumptive WT or heterozygous siblings.

Collectively, these results suggested a potential delay in specification, which initially led to 'trapping' of iridophore precursors in a premigratory position. An alternative interpretation is that a partially fate-restricted progenitor expressing *tfec* failed to undergo specification, thus becoming trapped in a premigratory position. Under the first hypothesis, the persistent reduction of iridophores during later specification and differentiation stages might indicate that a delay in iridoblast specification led to the onset of apoptosis. Under the second hypothesis, iridophore progenitors either underwent

specification towards an alternative lineage, or were lost by apoptosis. To test whether fate switch towards the melanophore lineage took place, the effects of loss of *foxd3* function on melanoblast specification were assessed using the marker *mitfa*.

At 18 hpf, there were no observable differences in the *mitfa* expression pattern between N = 54 embryos (2 independent experimental repeats; data not shown). At this stage, expression was consistently observed in a population of cells of the anterior dorsal trunk domain. The number of cells mildly varied between individuals, but this was attributed to small differences in the developmental stages of fixed embryos. At 24 hpf, a prominent reduction of *mitfa* positive migrating cells was observed in 12/56 assessed embryos (mutants according to Mendelian ratios; 2 independent experimental repeats; $0.5 < p\text{-value} < 0.7$; Fig. 5.2 A,B). Although cells residing along the dorsal trunk were not scored, it was likely that the missing migrating cells were restricted in that domain, indicating a defect either in melanoblast specification or migration. *mitfa* expression in cNCCs and NCC derivatives of the anterior trunk appeared grossly unaffected in presumed mutants compared to WT siblings (Fig. 5.2 A,B). A subset of these embryos was genotyped by HRMA and the genotypes were further validated by PCR-based genotyping (data not shown), which confirmed that the observed phenotype correlated with homozygosity of the *foxd3*^{sa20726} allele.

At 30 hpf, 15 embryos subjected to *in situ* hybridisation to detect *mitfa* expression were imaged and then genotyped (one experimental repeat). Only one embryo was confirmed to be homozygous mutant. In genotyped WT and heterozygous embryos, *mitfa* was expressed in melanoblasts and differentiated melanophores across the dorsal trunk and in the medial and lateral migratory pathways from the anterior trunk to the tail (Fig. 5.2 E). No observable differences in the number of *mitfa* positive cells along the anterior dorsal trunk region and in the migratory pathways were detected between the confirmed mutant embryo and WT siblings (Fig. 5.2, E,F). Towards the posterior trunk and tail, the premigratory domain of the mutant embryo appeared more strongly stained compared to WT embryos, and migratory cells were reduced, similarly to the phenotype observed at 24 hpf. These results suggested that by 30 hpf, melanophore progenitor numbers were restored, although further experiments are required to confirm that this altered phenotype manifests in 25% of embryos.

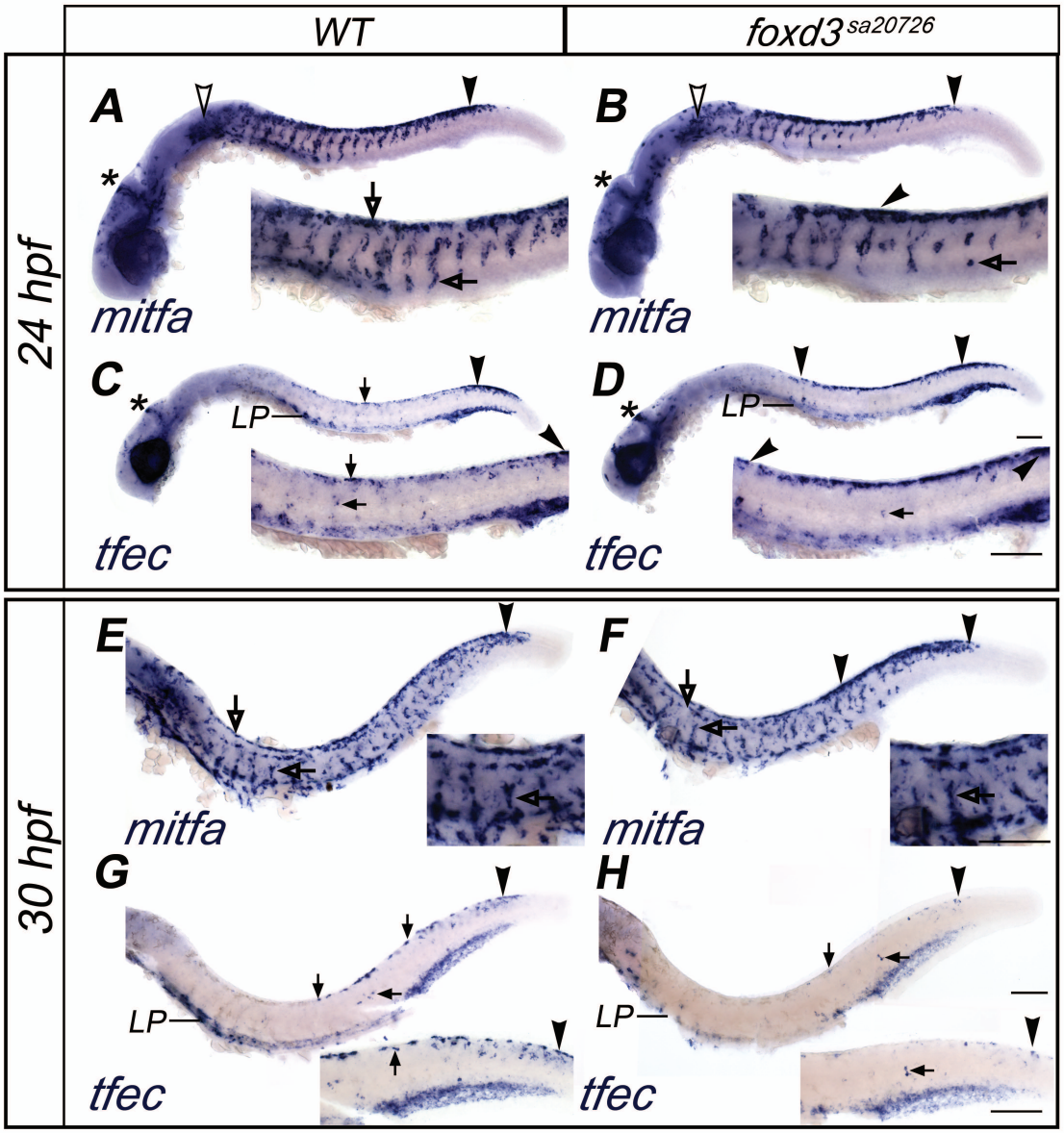


Figure 5.2. Chromatoblasts present with specification defects in *foxd3*^{sa20726} mutants, but the melanophore lineage recovers later in development. In WT embryos at 24 hpf (A), *mitfa* expression labels premigratory cells (arrowheads) and specified, migrating melanoblasts (insert, arrows). In confirmed *foxd3* mutants at the same stage (B) migrating cells (arrow) are decreased, with a number of progenitors potentially trapped within the premigratory domain (arrowhead). cNCCs (asterisks) and NCC derivatives of the anterior trunk (empty arrowheads) do not appear affected (A, B). At 24 hpf, WT *tfec* expression (C) is found in cNCCs (asterisk), unspecified NCCs along the posterior dorsal trunk (arrowheads), specified iridoblasts along the anterior dorsal trunk and in the migratory pathways (arrows) and in LP progenitors. In *foxd3* mutants (D), the premigratory NC domain is expanded anteriorly (arrowheads), cNCCs appear more prominently stained (asterisk) and LP progenitors are not detectable. Few migrating cells are still present (arrows). In the *foxd3* mutant embryo at 30 hpf (F), *mitfa* expression in anteriorly located melanoblasts (arrows) is not observably affected, compared to WT siblings (E). The mutant displays a lack of migrating cells towards the posterior, accompanied by potential expansion of the premigratory domain (arrowheads) (F). In phenotypically WT embryos at 30 hpf (G), *tfec* is normally expressed in iridoblasts along the dorsal trunk and the premigratory pathways (arrows), in the developing LPs and in premigratory NCCs in the tail (arrowheads). In presumed *foxd3* mutants (H), the number of stained iridoblasts (arrows) and of LP progenitors is reduced, while premigratory NCCs are still detectable (arrowheads). LP, lateral patches. Lateral views, oriented with the head to the left. Scale bars: 100 µm.

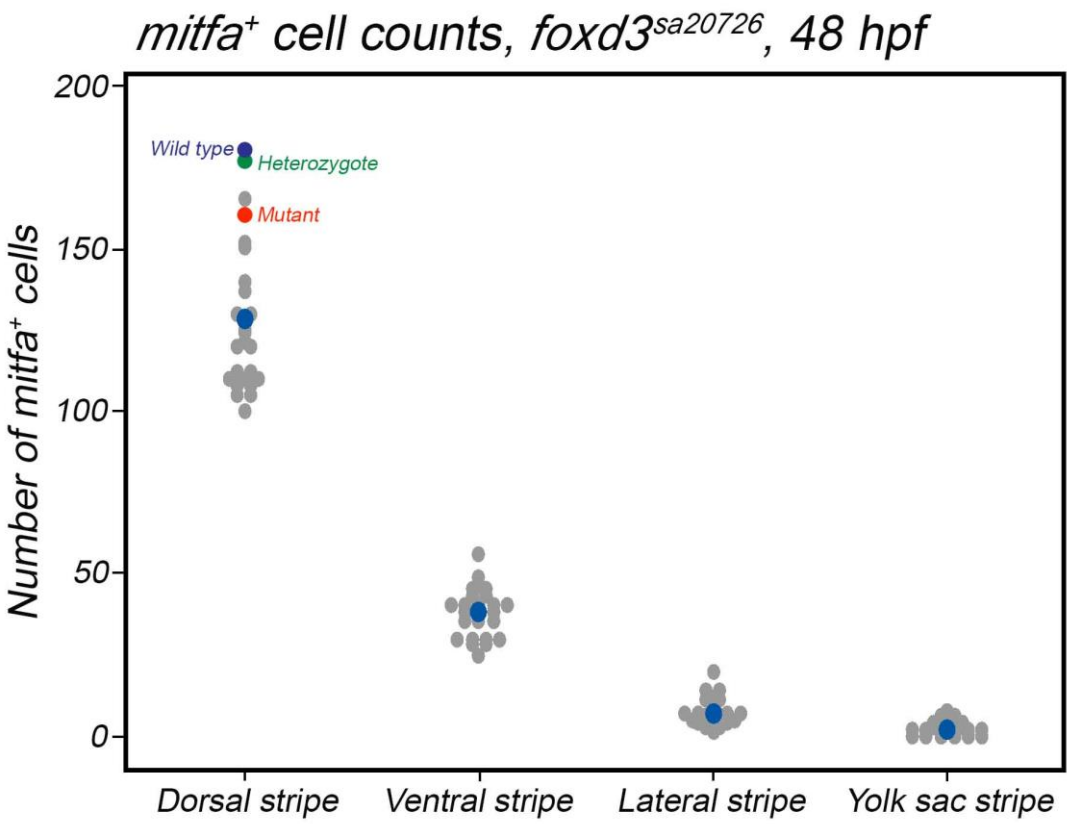


Figure 5.3. Numbers of *mitfa* positive cells do not detectably increase in *foxd3*^{sa20726} mutants. Counts of *mitfa* positive cells in N = 25 embryos comprising WT, heterozygous *foxd3*^{sa20726} carriers and homozygous mutants. Cell numbers along the ventral, lateral and yolk sac stripes cluster tightly. Molecular genotyping confirmed that the observed increase in *mitfa* positive cell numbers along the dorsal stripe in a subset of embryos (3 representatives indicated) is not a result of loss of *foxd3* function.

Finally, the experiment was performed in embryos fixed at 48 hpf (N = 25, one experimental repeat). *mitfa* positive cells were scored along the dorsal, ventral, lateral and yolk sac stripes (Fig. 5.3). While the number of cells along the ventral, lateral and yolk sac stripes clustered tightly, a relatively high degree of variability was observed in the number of *mitfa* positive cells along the dorsal stripe of individual embryos. The embryos were molecularly genotyped to determine whether individuals presenting with higher numbers of *mitfa* positive cells along the dorsal stripe corresponded to *foxd3*^{sa20726} homozygous mutants (data not shown). As a result, only one of the outliers was confirmed to be a homozygous mutant. The remaining 4 individuals with more than 150 *mitfa* positive cells along the dorsal stripe were heterozygotes (3/4) and WT (1/4) siblings. These data suggested that loss of *foxd3* function did not lead to an increase in *mitfa* positive cells owing to de-repression of *mitfa* in the iridophore lineage. Moreover, a fate switch from iridophores to melanophores was not supported.

Together, these data support a model according to which *foxd3* is important for the development of a subset of premigratory NCCs. The previously noted delay in melanoblast specification was confirmed, although the lineage appeared to gradually recover and melanoblast numbers in the trunk were restored by 30 hpf. Moreover, a defect in iridophore specification was reported, from which the lineage failed to recover. Although the nature of the defect remains unclear, a fate switch from the iridophore to the melanophore lineage was not supported.

5.3. *foxd3* is expressed in a subset of iridoblasts and iridophores at low levels

5.3.1. *foxd3* expression is not detectable by *in situ* hybridisation

To further investigate the role of *foxd3* in the iridophore lineage, *foxd3* expression in mature iridophores was assessed by conventional *in situ* hybridisation. Iridophores of live WT embryos (WIK background) were imaged at 2 dpf (Fig. 5.4 A,B). The embryos were then fixed and *in situ* hybridisation was performed to detect *foxd3* transcripts (Fig. 5.4 C,D). Expression was not detectable in iridophore locations along the dorsal trunk or on the lateral patches of the embryos. However, prominent expression was detectable in DRG positions and oligodendrocytes, which are cell types known to express *foxd3* (Fig. 5.4 A-D).

5.3.2. A subset of iridoblasts and iridophores express *foxd3*

RNAscope was employed to more extensively evaluate *foxd3* expression dynamics over time in the iridophore lineage. Co-expression of the iridoblast and iridophore marker *ltk* and of *foxd3* was assessed in embryos fixed at 24 hpf, 30 hpf, 36 hpf and 48 hpf.

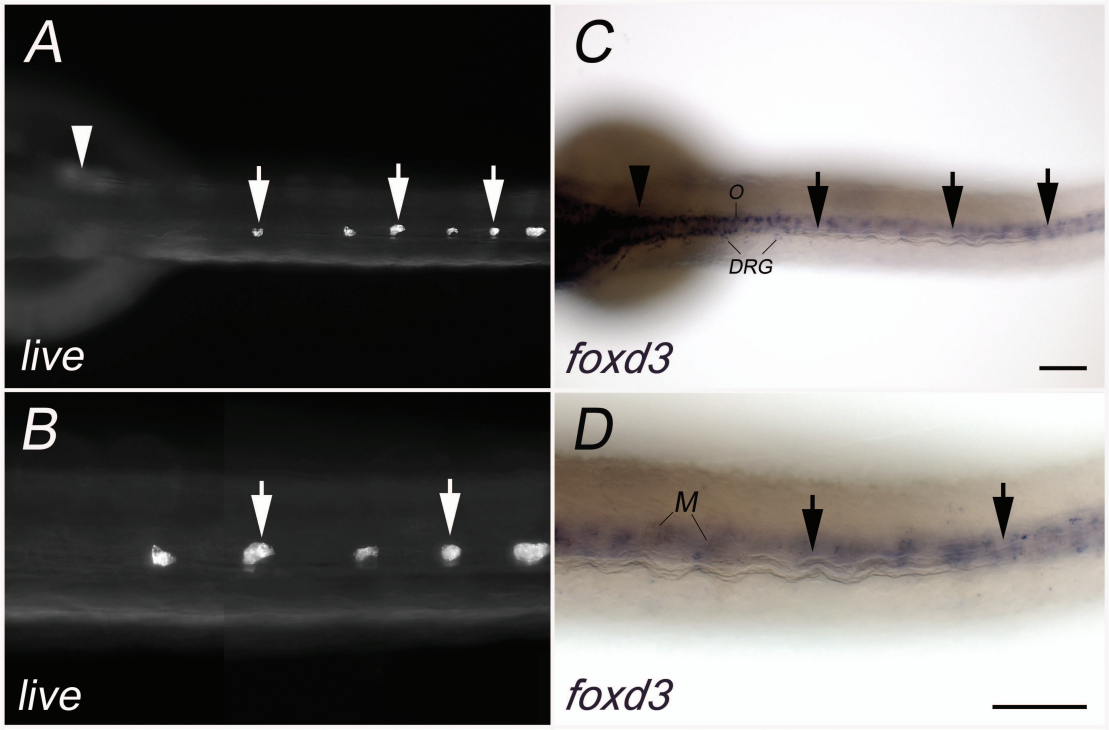


Figure 5.4. *foxd3* transcript is not detectable in mature iridophores by chromogenic in situ hybridisation at 2 dpf. Live imaging of a PTU-treated embryo under incident light (A, B) reveals iridophore locations along the dorsal trunk (arrows) and on the lateral patches (arrowhead). Performing in situ hybridisation on the same embryo (C, D) indicates lack of *foxd3* expression in relevant iridophore locations along the trunk (arrows) and on the lateral patches (arrowhead). *foxd3* is detectable in oligodendrocytes in the vicinity of the neural tube and at DRG positions. o, oligodendrocytes; DRG, dorsal root ganglia; M, melanocytes. Dorsal views, oriented with the head towards the left. Scale bars: 100 µm.

To assess co-expression, all *ltk* positive cells were identified along the dorsal trunk of each embryo. A cell was considered positive for gene expression when at least two strong fluorescent spots were detected adjacent to the nucleus, which was visualised using DAPI counterstaining. Iridoblasts and iridophores only of the dorsal trunk were considered for two main reasons: firstly, prominent *foxd3* expression in somites could be misinterpreted for signal in medially migrating iridoblasts. Second, background signal within the notochord was commonly observed in RNAscope samples and could obscure signal in migrating cells. Following identification of all dorsal trunk *ltk* expressing cells, each was scored for *foxd3* expression. As in previously described studies, co-expression of two genes was concluded if their respective transcripts visibly overlapped and were in close proximity to the periphery of a DAPI-labelled nucleus. This convention prevented interference from cytoplasms of adjacent cells likely expressing one of the genes. Notably, it has been previously reported that individual spots of fluorescence observable in RNAscope assays represent single mRNA molecules (F. Wang et al. 2012). Thus, quantitative assessment of relative expression levels was conducted by scoring the number and intensity of fluorescent spots surrounding distinct cell nuclei. In order for quantitation to become absolute, it is required to measure fluorescence intensity from confocal microscope images, using specialised software (Gross-Thebing et al. 2014).

Surprisingly, results obtained using this methodology indicated that approximately 50% of *ltk* positive specified iridoblasts and mature iridophores co-expressed *foxd3* across all developmental stages investigated (Fig. 5.5; Fig. 5.6 A). At 24 hpf, N = 128 *ltk* positive cells of the dorsal trunk of 5 embryos were scored for *foxd3* expression, which was observed in 54.9% of cells (Fig. 5.5 D; Fig. 5.6 A). Low level *ltk* expression, represented by a small number of easily resolved spots localised around distinct nuclei, was detectable in premigratory NCCs (Fig. 5.5 A, inset 2). These might represent either multipotent progenitors or oligopotent cells at the early stages of fate restriction. Intense labelling, indicating relatively high expression levels was seen in a subset of cells, likely corresponding to specified iridoblasts (Fig. 5.5 A). Prominent *foxd3* expression at this stage was observed in a number of NCCs and neural derivatives within the neural tube (Fig. 5.5 B), as well as in somites (data not shown).

At 30 hpf, *ltk* expression prominently labelled specified iridoblasts along the dorsal trunk and the migratory pathways (Fig. 5.5, E). *foxd3* was strongly expressed in somites and in PNS components of the neural tube, while expression along the dorsal trunk had significantly subsided (Fig. 5.5 F). Co-expression of *ltk* and *foxd3* was detected in 63.8% of iridoblasts (N = 183 cells from 4 embryos) residing along the dorsal trunk (Fig. 5.5 H; Fig. 5.6 A).

At 36 hpf, *ltk* expression was detectable in iridoblasts along the dorsal trunk, while *foxd3* expression strongly labelled somites and neural derivatives (Fig. 5.5 I,J). At this stage, co-expression was detectable in 43.7% of *ltk* positive iridoblasts (N = 219 cells from 7 embryos) along the dorsal trunk (Fig. 5.5 L; Fig. 5.6 A). Finally, at 48 hpf, N = 126 *ltk* positive iridophores from 3 embryos were scored and co-expression was detected in 51.6% (Fig. 5.5 P; Fig. 5.6). At this stage, *foxd3* was prominently expressed in neural derivatives located within the neural tube (Fig. 5.5 M,N) and in the somites (data not shown).

At 24 hpf, iridoblasts expressed relatively higher levels of *ltk* than NCCs, as was indicated by the number and intensity of fluorescent spots in these cells. In order to test whether increased levels of *ltk* related to increased or decreased levels of *foxd3,* relative transcript levels both of *ltk* and of *foxd3* on each *ltk* positive cell were determined by scoring the number of spots surrounding individual nuclei. These data were used to perform linear regression analyses (refer to section 2.8.3). Results indicated weak but statistically significant negative correlation (r^2 = 4.6%, p-value = 0.013) between *ltk* and *foxd3* expression levels (Fig. 5.6 B). It was therefore suggested that *foxd3* was not upregulated in an *ltk*-dependent manner within the iridophore lineage but potentially functioned in NCC derivatives undergoing specification, which were in the process of downregulating *ltk*.

In conclusion, expression data in WT embryos suggested that *foxd3* was maintained in a subset of iridophores throughout development. This expression was relatively weak and could not be detected by conventional *in situ* hybridisation. Expression of *sox10*, however, was retained in every cell in the iridophore lineage (refer to section 4.4.3), hence the repressor of *mitfa* was required to be expressed in a similar pattern. These results argued against *foxd3* fulfiling the role of factor R in the iridophore lineage.

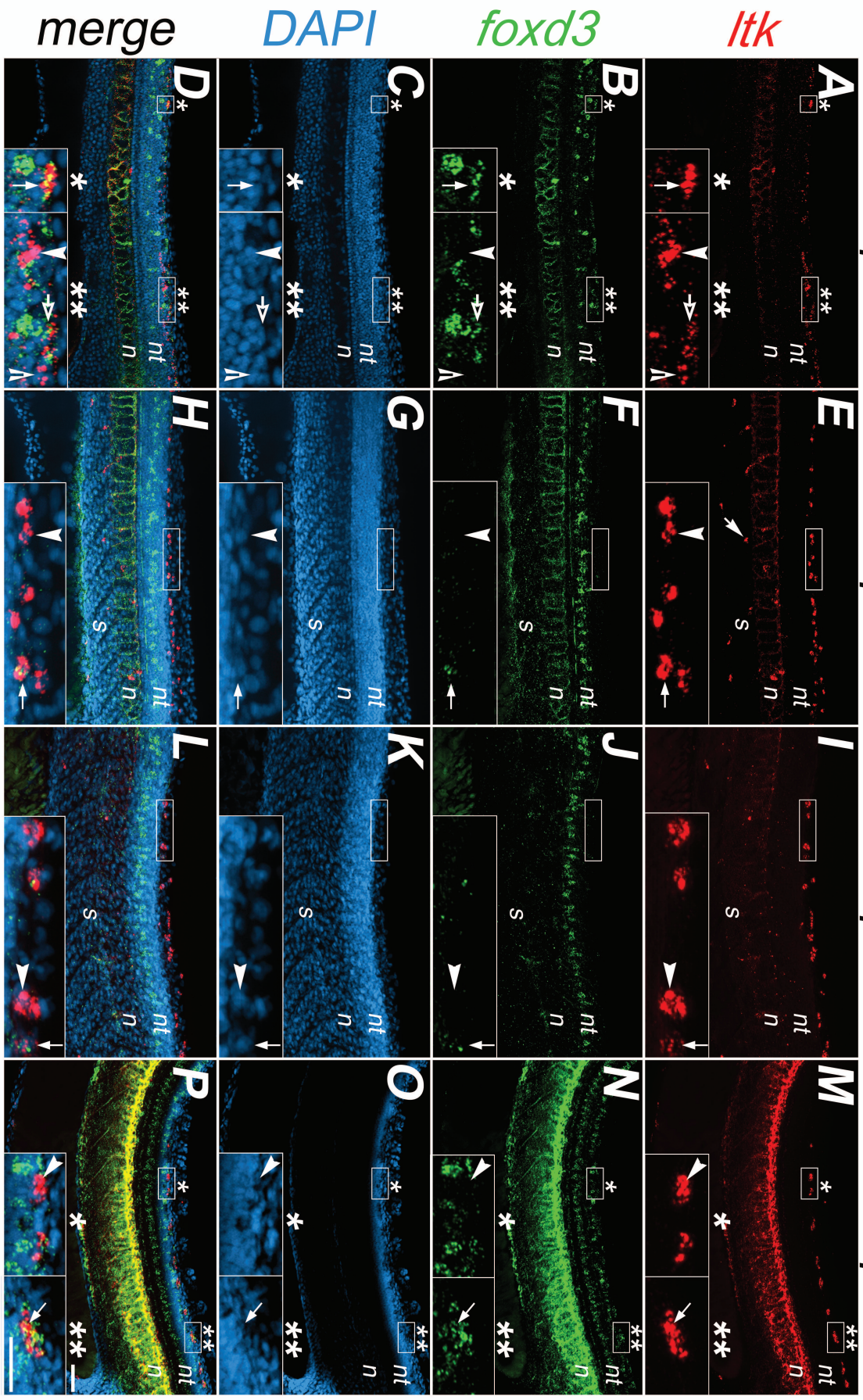

merge
DAPI
foxd3
ltk
24 hpf
30 hpf
36 hpf
48 hpf
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
nt
n
s

Fig. 5.5

Figure 5.5. *foxd3* is expressed in approximately 50% of specified iridoblasts and mature iridophores across developmental stages. RNAscope results indicate co-expression of *ltk* and *foxd3* (arrows) in a subset of iridoblasts and mature iridophores throughout development, while approximately 50% of *ltk* positive cells are consistently *foxd3* negative (arrowheads). At 24 hpf (A-D), (A) low level *ltk* expression labels *foxd3* negative (split arrowhead) and *foxd3* positive (empty arrow) NCCs (inset 2). Increased level of *ltk* labels *foxd3* positive (full arrow) and negative (full arrowhead) iridoblasts (insets 1 and 2). *foxd3* expression (B) is detectable in the vicinity of the neural tube. At 30 hpf (E-H), a subset of *ltk* positive iridoblasts co-express *foxd3* (inset, arrows), while *foxd3* (F) is expressed primarily in somites and along the neural tube. At 36 hpf (I-L), a subset of iridoblasts (inset, arrows) co-express *ltk* and *foxd3*. The latter is detectable in somites and neural derivatives (J). By 48 hpf (M-P), *ltk* labels iridophores (M) and *foxd3* is mainly expressed in oligodendrocytes along the neural tube and in other neural derivatives (N). A *foxd3* negative (inset 1, arrowheads) and a *foxd3* positive (inset 2, arrows) population of iridophores is noticeable. In all stages, the notochord presents with background signal from both channels. nt, neural tube; n, notochord; s, somites. Lateral views, oriented with the head to the left. Scale bar: 50 μm. Inset scale bar: 25 μm.

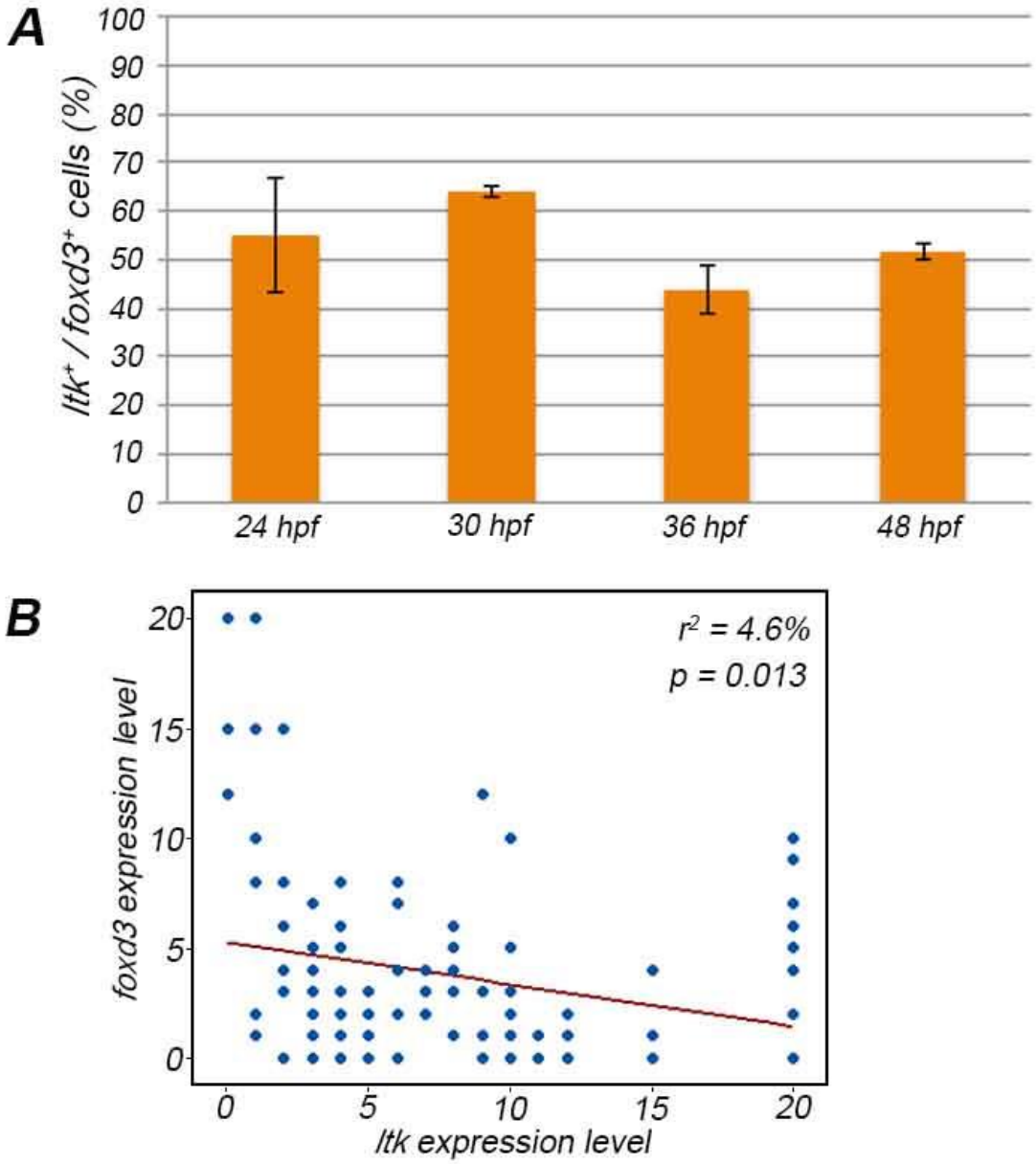


Figure 5.6. *foxd3* expression is maintained in a sub-population of cells of the iridophore lineage. (A) Percentages of *ltk*$^+$/*foxd3*$^+$ cells along the dorsal trunk at 24 hpf (N = 128 *ltk*$^+$ cells scored), at 30 hpf (N = 183), at 36 hpf (N = 219) and at 48 hpf (N = 126). Error bars indicate the standard error of the mean across biological replicates (number of embryos analysed per stage). (B) Linear regression analysis results from semi-quantitative analysis of gene expression at 24 hpf. A very weak (r^2 = 4.6%) but statistically significant (p-value < 0.05) negative correlation was observed between increasing level of *ltk* expression and the corresponding level of *foxd3* expression in individual cells. Expression levels are derived from the number and intensity of fluorescent spots representing each transcript, which surround distinct nuclei.

5.4. Investigating the role of *id2a* in repressing *mitfa*

5.4.1. *id2a* is strongly co-expressed with iridophore markers

Next, the known repressor *id2a* was investigated as a candidate for the role of the factor R in the iridophore lineage. To identify whether the *id2a* expression pattern matched expectations for the *mitfa* repressor, RNAscope studies were performed using *ltk* and *id2a* probes simultaneously. In line with previous reports (F. R. Rodrigues, PhD thesis), results indicated that *id2a* was expressed in the majority, if not all, of premigratory tNCCs at 24 hpf (data not shown). By 36 hpf *id2a* was expressed strongly in all specified iridoblasts (Fig. 5.7), matching the key requirement for factor R expression dynamics.

5.4.2. Generation of *id2a* mutants

If *id2a* indeed functions to repress *mitfa* in the iridophore lineage, it was predicted that in *id2a* mutants *mitfa* expression would be upregulated. To directly test this prediction, *id2a* mutants were generated using CRISPR/Cas9 technology. To generate mutants, two guide RNA sequences targeting different regions of *id2a* (refer to Appendix IV) were tested. These are referred to as *id2a* CRISPR1 or CRISPR2. Injections were carried out according to guidelines outlined in section 2.2.2 and injected embryos were examined under incident light to assess possible phenotypic effects on the iridophore lineage. There was no observable loss of iridophores in injected individuals (F0 generation), therefore molecular techniques were employed to screen for mutagenesis.

To estimate the cutting efficiency for each guide RNA, 7-8 embryos injected either with *id2a* CRISPR1 or CRISPR2 were pooled, their genomic DNA was extracted and the region surrounding the target sequence was amplified by PCR (refer to section 2.2.2; Appendix II). Single PCR amplicons were cloned into suitable vectors (refer to section 2.4). Individual clones for each guide were sequenced and sequences were aligned to detect alterations (refer to section 2.3.6). None of the amplicons derived from CRISPR2-injected fish represented mutagenized sequences (data not shown), but 6 out of 16 sequences obtained from CRISPR1 injections showed distinct alterations (Fig. 5.8 A). Identified mutations were mostly deletions, but nucleotide insertions were also detected. In 2 out of 6 cases, an insertion and a deletion had occurred simultaneously (Fig. 5.8 A), an effect which has been previously described as a result of CRISPR/Cas9 mutagenesis (Hwang et al. 2013). CRISPR1-injected embryos (WIK background) were raised to adulthood.

To screen for adults transmitting mutagenised *id2a* alleles through their germline, individuals were first incrossed to perform preliminary screening (Fig. 5.8 B). Individual fish from pairs transmitting mutant alleles to their progeny were outcrossed to WT adults (WIK background) to identify founders (Fig. 5.8 C-E) and to raise the F1 generation. For

screening purposes genomic DNA was extracted from 16 single embryos derived from each cross and HRMA was performed (refer to section 2.3.7) to identify carriers of mutant alleles. 16 known WT embryos were used in each assay as reference samples (data not shown). The HRMA software analysed melting temperature variations between amplicons derived from reference WT samples and from unknown samples to identify altered sequences. As a result, 3 adults transmitting altered sequences to their progeny were identified. Such sequences were represented by 'variant 1' downward deflecting difference curves (Fig. 5.8 C, D, E). Incrossed adults that did not transmit mutant alleles to their offspring (Fig. 5.8. B) were euthanised. Siblings of screened embryos bearing *id2a* mutant alleles were raised to adulthood. As discussed in section 4.2.1, further screening is required to generate clean lines of single mutant alleles.

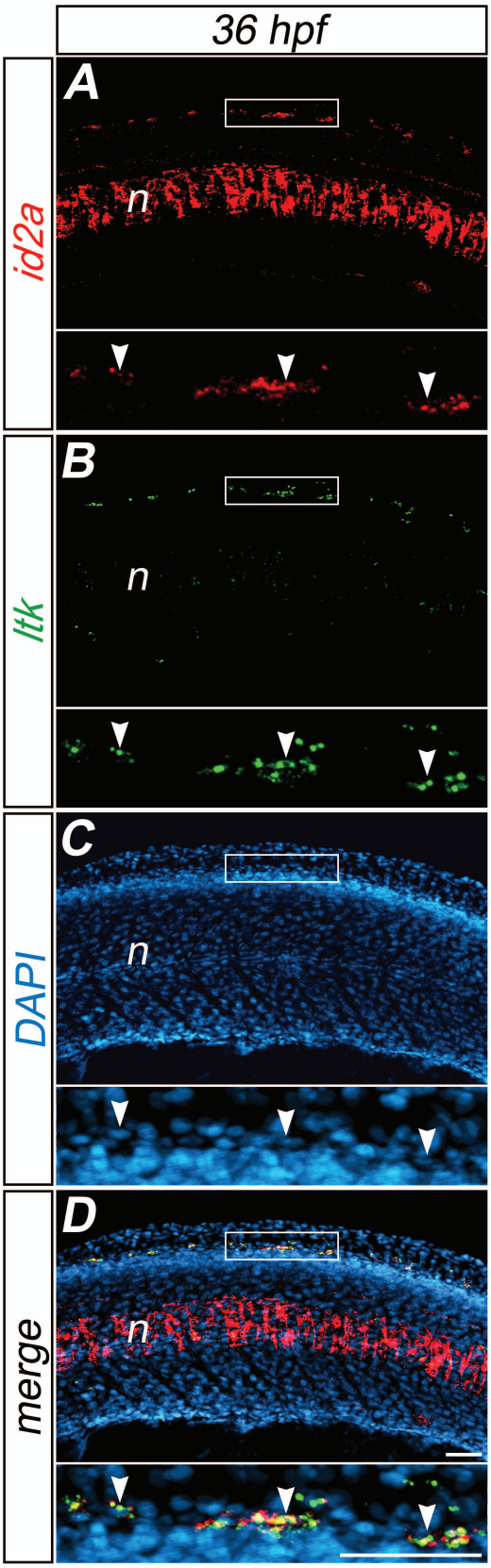


Figure. 5.7. The *id2a* expression pattern fulfils factor R requirements. RNAscope results at 36 hpf indicate co-expression of *ltk* and *id2a* in all iridoblasts located along the dorsal trunk (inset, arrowheads). High levels of background fluorescence are detectable along the notochord in the *id2a* channel (A, D). n, notochord. Lateral views, oriented with the head to the left. Scale bars: 50 µm.

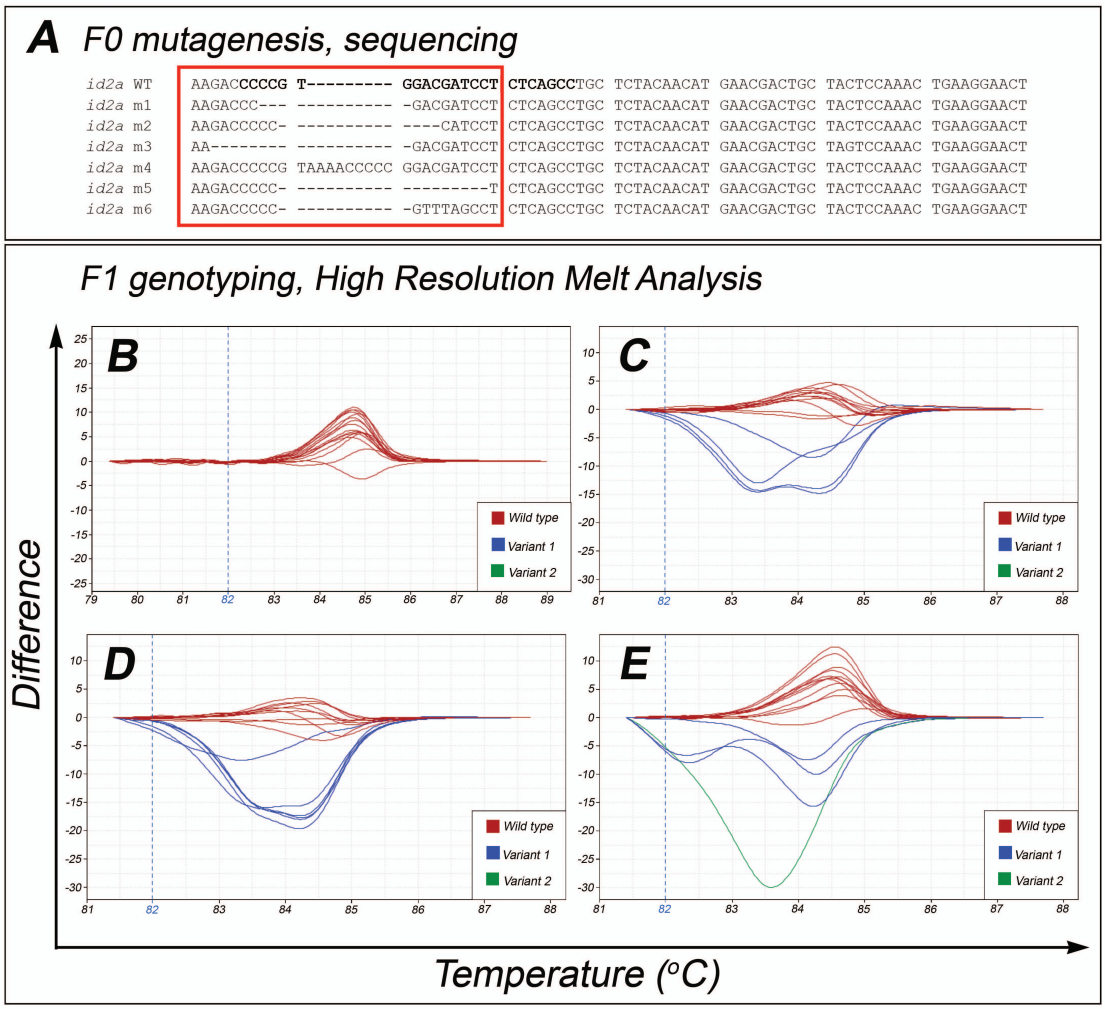


Figure 5.8. Molecular identification of CRISPR/Cas9 mutagenesis in the id2a locus. (A) Sequencing of individual clones containing *id2a* WT and mutant (m) alleles from injected fish. The extent of insertions (*id2a* m2, m4 and m6) and deletions (*id2a* m1, m2, m3, m5 and m6) is indicated by the red box. The CRISPR target sequence is shown in bold in the WT sequence. HRMA-based identification of (C-E) F0 founders transmitting *id2a* mutant alleles to the F1 generation and (B) of injected adults that only transmit WT alleles. Embryos genotyped in (B) were derived from an incross of F0 adults, while embryos represented in (C-E) were derived from outcrosses of F0 to WT fish (WIK background). Variant 2 (E) does not represent mutagenesis, but results from low amplification efficiency in the respective sample.

5.5. Discussion

5.5.1 Testing candidates for factor R

This chapter focused on the experimental identification of the iridophore-specific repressor of *mitfa*, factor R, which was introduced in model D. The key requirements for factor R were the following:

1. Factor R should be expressed in all cells of the iridophore lineage in a manner similar to that demonstrated for *sox10*.
2. Upon loss of factor R function, *mitfa* should be de-repressed in the iridophore lineage.

The first candidate tested for the role of factor R was *foxd3*, which codes for a TF consisting of 371 amino acid residues. The major functional domains of Foxd3 comprise an N-terminal acidic domain, responsible for transactivation, and a widely conserved forkhead domain, which mediates DNA binding (Wijchers, Burbach, and Smidt 2006). The forkhead domain, which is a type of winged-helix domain, comprises three consecutive α-helices and associated β-strands (Wijchers, Burbach, and Smidt 2006). Structural studies of the forkhead/DNA complex have shown that DNA recognition occurs through the third consecutive α-helix (Clark et al. 1993), while the two other helices have been suggested to function to stabilise the complex and to confer DNA-binding specificity (Wijchers, Burbach, and Smidt 2006).

Foxd3 function has been reported in several contexts, including ESC and NCC establishment and maintenance (Dottori et al. 2001; Hanna et al. 2002; Teng et al. 2008). Interestingly, Foxd3 has been suggested to act both as an activator and as a repressor of transcription (Thomas and Erickson 2009; Powell et al. 2013; Sweet et al. 2016), with a newly appreciated potential role as a pioneer TF (refer to section 1.5.4). In regards to pigment cell development, Foxd3 has been previously suggested to repress Mitf in several contexts and model organisms (Kos et al. 2001; Ignatius et al. 2008; Thomas and Erickson 2009; Curran, Raible, and Lister 2009; Curran et al. 2010). In particular, studies using zebrafish as a model indicated that *foxd3*-dependent repression of *mitfa* was important in the melanoiridoblast fate decision process, allowing iridophores to become specified (Curran et al. 2010).

Taking into account this body of evidence supporting the role of *foxd3* as an inhibitor of *mitfa* in the iridophore lineage, *foxd3* was first evaluated for the role of factor R in model D. Intriguingly, results from this work contradicted the currently accepted role of *foxd3* as the repressor of *mitfa* during iridophore development, as neither the first nor the second requirement were met. Specifically, quantitative studies suggested an inverse correlation between *ltk* and *foxd3* expression levels during early stages of specification, indicating

that *foxd3* expression was likely not activated in iridoblasts. Importantly, *foxd3* expression was detected in approximately 50% of specified iridoblasts and differentiated iridophores, in keeping with previously published data where only a subset of mature iridophores was shown to express *foxd3* (James A Lister et al. 2006). This result indicated a role for the gene in a putative iridophore subpopulation.

In contrast, *sox10* expression was demonstrated in the majority, if not all, of iridoblasts and mature iridophores. Since *mitfa* is a direct target of *sox10* (Elworthy et al. 2003), it follows that in the absence of a repressor it would become upregulated in the iridophore lineage. However, both data from the literature (J A Lister et al. 1999) and results presented in a previous chapter argue against *mitfa* maintenance in the iridophore lineage throughout its development. Therefore, based on the first criterion, *foxd3* was concluded to be insufficient as a repressor of *mitfa* expression in this context.

With regards to the second requirement, loss of *foxd3* function did not result in supernumerary *mitfa* positive cells. In particular, melanoblast specification was delayed suggesting that, at early stages, *mitfa* activation in the lineage might in fact require *foxd3* function. Moreover, the number of cells expressing *mitfa* at differentiation stages did not change between *foxd3* mutants and WT siblings. It was concluded that lack of *mitfa* inhibition by *foxd3* was not responsible for the loss of iridophores in *foxd3* mutants, therefore contradicting the current view regarding the role of *foxd3* in the melanoiridoblast (Curran et al. 2010), and that *mitfa* was not de-repressed in iridophores upon loss of *foxd3* function. Collectively, these observations suggested that *foxd3* might be able to repress *mitfa* in certain contexts, but that was not likely to be the case in iridophores.

Having excluded *foxd3*, *id2a*, was assessed for the role of factor R in the GRN. Id family proteins function to sequester bHLH TFs, such as *mitfa*, by forming heterodimers via their own HLH domain. However, these proteins lack the basic (b) domain, which is crucial for interaction with the DNA (Ling, Kang, and Sun 2014). Id family TFs have not been widely studied, especially in the context of pigment cell development. A previous PhD student in our research group identified a role for *id2a* in iridophore development (F. Rodrigues, PhD thesis). Specifically, *id2a* was shown to be expressed in mature iridophores, while morpholino-mediated knock-down of the gene resulted in complete lack of iridophores. Taking the above into account, *id2a* was considered to be a good candidate for the role of the iridophore-specific *mitfa* repressor in the iridophore GRN.

Data from the present study confirmed that the *id2a* expression pattern fulfils the first requirement for factor R, namely prominent expression was shown in all cells of the iridophore lineage. To enable testing of the effects of loss of *id2a* function in the iridophore lineage, *id2a* mutant lines were required, although such lines were not

available at the time. To generate *id2a* mutants, the CRISPR/Cas9 system for site-directed mutagenesis was employed. To date, F0 founders transmitting *id2a* mutant alleles through their germlines have been identified and their progeny has been raised to adulthood. Before these mutants can be thoroughly analysed, it is important to identify at least two or three F1 founders and use them to generate pure mutant lines.

Upon obtaining *id2a* mutant lines, the second requirement will be addressed by directly testing the expression pattern of *mitfa* in mutant embryos compared to their WT siblings. If the hypothesis that *id2a* represses *mitfa* in the iridophore lineage is valid, then de-repression of *mitfa* in these cells will become obvious, at least at differentiation stages. Moreover, it is conceivable that in *id2a* mutants the onset of *mitfa* expression in the premigratory NC domain will be earlier than normal, since *id2a* is known to be expressed in the majority, if not all, of tNCCs in premigratory positions between prior to 24 hpf (F. Rodrigues, PhD thesis), at stages when *mitfa* is not detectably expressed (J A Lister et al. 1999).

To further understand the potential role of *id2a* in *mitfa* repression, the expression pattern of direct targets of *mitfa*, such as the melanophore differentiation gene, *dct*, should be assessed. Since *id2a* acts by sequestering TFs, thus preventing them from binding to DNA, rather than by directly interfering with their expression, it is likely that *mitfa* repression by *id2a* relies upon inhibition of *mitfa*-dependent gene activation. Since upregulation of *mitfa* itself has been demonstrated to occur in a *mitfa*-dependent manner (Greenhill et al. 2011), it is postulated that inhibiting *mitfa* function leads to downregulation of *mitfa* expression. Therefore, lack of *mitfa* transcript in developing iridophores could be justified based on its repression by *id2a*, an interaction which should be directly tested using *id2a* mutants.

5.5.2. Characterisation of a novel *foxd3*^{sa20726} mutant allele

A controversy with regard to the role of *foxd3* in pigment cell development has been established in the literature. Using different *foxd3* mutant lines, several studies have indicated that melanophore specification from the NC is initially delayed, although melanophore numbers are gradually restored (James A Lister et al. 2006; Montero-Balaguer et al. 2006; Stewart et al. 2006). However, the numbers of mature iridophores were shown to be reduced, with the effect being more prominent in the anterior dorsal stripe and absent in the ventral stripe and the eye of the embryo (James A Lister et al. 2006; Montero-Balaguer et al. 2006; Stewart et al. 2006). Furthermore, a severe jaw malformation phenotype was identified as a result of loss of *foxd3* function (James A Lister et al. 2006; Montero-Balaguer et al. 2006; Stewart et al. 2006). Results from these studies thus appear to be consistent in that the melanophore lineage is not affected by lack of functional *foxd3*, whereas the iridophore lineage is significantly, but not totally,

reduced. This view was subsequently challenged by data suggesting that *mitfa* expression is expanded upon loss of *foxd3* function, and that *foxd3* physically binds the *mitfa* promoter to repress *mitfa* activation (Curran, Raible, and Lister 2009). According to a more recent report, mutants for another zebrafish *foxd3* allele presented with a striking expansion of melanophores, which was not previously reported in the literature (Hochgreb-Hägele and Bronner 2013).

Inconsistencies were thus identifiable in the published literature, likely arising from the fact that different mutant lines or knockdown approaches were used in each study (refer to section 5.1). Briefly, it was observed that in all of the described mutant lines the N-terminus of *foxd3*, containing the acidic transactivation domain, remained intact. In the case of the *sym1* mutant line, in addition to the acidic domain, the first and third α-helices were also not disrupted (Stewart et al. 2006). Moreover, the increase in melanophore number phenotype associated with the gt(*foxd3*-cherry)^{c110aR} line suggests that insertion of the transgene resulted in non-physiological functions. Therefore, the need for a further mutant line, in which a larger portion, preferably all, of the protein is eliminated, arose in order to obtain a better understanding of *foxd3* function in pigment cell development.

Here, the currently uncharacterised mutant line, *foxd3*^{sa20726}, which features a premature stop codon at the 40th of 371 amino acid residues of Foxd3, was used in an attempt to reconcile this controversy. In accordance with published research, the previously reported delay in melanogenesis, as well as the reduction in the numbers of differentiated iridophores along the dorsal, but not the ventral stripe were confirmed. Intriguingly, the jaw malformation phenotype, which was identified in all of the previously described mutant lines, was only partially penetrant and the ectopic melanophore phenotype reported in the transgenic gt(*foxd3*-cherry)^{c110aR} line (Hochgreb-Hägele and Bronner 2013) was not reproduced using this allele. It is plausible that elimination of all the major domains of the protein, including the acidic transactivation domain, which stretches from the 21st to at least the 60th residue of the protein according to the sequence provided in the Ensembl database, was responsible for the differences observed in the jaw phenotype, although this hypothesis requires further testing. The lack of an increase in melanophore numbers presented here was consistent with previously published literature (James A Lister et al. 2006; Montero-Balaguer et al. 2006; Stewart et al. 2006), thus suggesting that the transgenic gt(*foxd3*-cherry)^{c110aR} line presents with neomorphic effects.

5.5.3. The role of *foxd3* in the iridophore lineage

Furthermore, results presented in this chapter addressed the events taking place in the early stages of iridophore development, with the view to identify the function of *foxd3* in the specification of this chromatophore lineage. The identification of *tfec*-expressing cells

trapped in the premigratory NC domain, which however appear less in number than those in *sox10* mutants (Lopes et al. 2008; Chapter 3), suggested either a delay in fate choice, similar to that observed in the melanophore lineage, or failure of specification of an NC-derived precursor which is at least responsible for generating iridophores.

The observed late onset of melanophore migration, combined with the absence of sustained defects in this lineage, supported the notion that NC-derived progenitors did not fail to become specified. Instead, it was considered plausible that specification was temporally delayed or, alternatively, that specified melanoblasts were transiently trapped in a dorsal position unable to migrate. Furthermore, in the iridophore lineage, normal numbers and morphologies of cells located along the ventral stripe in *foxd3* mutants reinforced the hypothesis that specification of the lineage had occurred normally, although less cells may have formed or survived.

Under the hypothesis that a delay in specification resulted from loss of *foxd3* function, the trapped *tfec* positive cells along the premigratory NC domain, which were present in *foxd3* mutants but not in WT siblings, could be interpreted as transiently persisting multipotent NCCs expressing *tfec*. The subsequent permanent partial loss of iridophores, as opposed to recovery similarly to the melanophore lineage, could be due to failure of iridoblasts to overcome this delay in specification. Alternatively, *foxd3* might be required for iridophore survival, which is however unlikely, as *foxd3* was shown to only be expressed in 50% of iridoblasts.

An alternative interpretation regarding the presence of trapped progenitors in *foxd3* mutants is that, like in *sox10* mutants, a particular type of NC-derived progenitor failed to become specified, to maintain or downregulate *tfec* and to start migrating. Interestingly, close observation of migrating *mitfa* positive cells in *foxd3* mutants indicated that a significant number of medially migrating melanoblasts along the trunk were missing, even though migration along the lateral pathway, which normally occurs approximately 3 hours later than in the medial pathway, had started. Although this could be explained by delay of a particular type of melanoblast to become specified, it may also indicate failure in specification of a precursor that is able to generate both melanoblasts and iridoblasts. Under this hypothesis, normal numbers of melanophores could potentially be restored through proliferation of migrating progenitor cells, while the persisting mature iridophores observed in *foxd3* mutants might be generated from a different type of precursor. This model of melanophore and iridophore specification fits

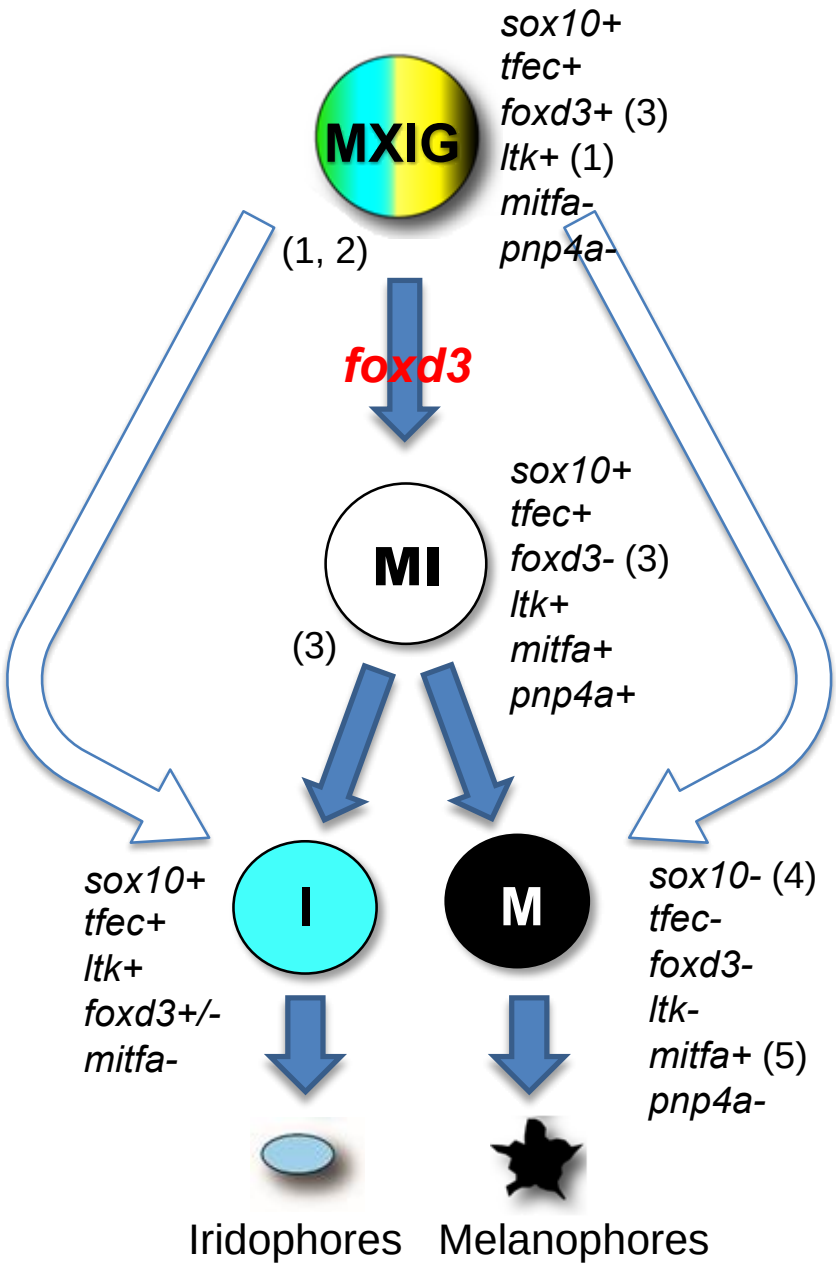


Figure 5.9. The working model for chromatophore development revisited. Previous data indicated that the chromatoglioblast (MXIG) (**1**, Lopes et al. 2008; **2**, M. Nikaido, unpublished data) generates different types of bipotent progenitors, one of which is the melanoiridoblast (MI) (**3**, Curran et al. 2010). This work focused on the identification of likely gene expression signatures for each intermediate cell type, which are indicated in the diagram. In particular, loss of *foxd3* function phenotypes, suggested that *foxd3* plays a role in the specification of the melanoiridoblast (indicated in red), as a subset of iridoblasts (I) and melanoblasts (M) were affected. The unaffected cells presumably have different origins, consistently with the previously established model (Curran et al. 2010). Other references: **4**, Greenhill et al. 2011; **5**, Lister et al. 1999.

the previously described model for the melanoiridoblast, under which iridophores and melanophores originate from a common NC-derived progenitor, but also from their own distinct progenitor populations (Fig. 5.9) (Curran et al. 2010).

The two models are, in some ways similar, as the first predicts a transient and the second a permanent defect in specification, thus providing experimental evidence to distinguish between the two hypotheses could prove challenging. A way of determining the developmental potency of cells trapped in the premigratory position in *foxd3* mutants would be by conducting co-expression studies to assess presence of markers for different cell types, specifically melanophore and iridophore markers. Furthermore, assays to determine apoptosis over a time course would provide evidence regarding the fate of trapped cells that failed to generate iridophores. Finally, co-labelling for *mitfa* expression and phosphohistone H3 using RNAscope would indicate whether migrating *mitfa* positive cells actively proliferate, thus whether this could result in restoration of normal numbers in the lineage.

5.5.4. Summary

In summary, this chapter focused on identification of a repressive factor R, which functions to downregulate *mitfa* in an iridophore specific manner. The previously established *mitfa* repressor failed to fulfil key requirements for factor R and was, thus, excluded. The process of evaluated the known bHLH repressor, *id2a*, for the role of factor R has been initiated. In the course of investigating the role of *foxd3* in chromatophore development, a novel *foxd3* mutant line was characterised and certain pigment cell specification phenotypes were identified. These could be interpreted as a delay in specification, which has previously been reported in the literature (James A Lister et al. 2006; Montero-Balaguer et al. 2006; Stewart et al. 2006), or as indicative of a common progenitor which requires *foxd3* to become specified, similarly to what has been previously suggested for *sox10* (Lopes et al. 2008).

6

General Discussion

6.1. Background

6.1.1. NC development is governed by complex GRNs

All cellular processes are underlain by specialised GRNs. During embryogenesis, highly dynamic and tightly regulated GRNs function to drive cells through consecutive transitions (Levine and Davidson 2005). The vertebrate NC has been recognised as an attractive system to study dynamic developmental transitions, including stem cell maintenance and proliferation, delamination, migration and, importantly, diversification of an astonishingly wide array of derivatives (Stemple and Anderson 1992; Brian K Hall 2008; Trainor 2014). Over the course of studying the developmental origins of NCCs and their derivatives, important genes governing the processes of interest have been identified (Sauka-Spengler and Bronner-Fraser 2008; Groves and LaBonne 2014). In recent years, several endeavours aimed to bring the identified genes and their interactions together to construct a series of GRNs that function during distinct stages of NC development (Baxter et al. 2010; Paola Betancur, Bronner-Fraser, and Sauka-Spengler 2010; Greenhill et al. 2011; Theveneau and Mayor 2012; M. Simões-Costa et al. 2015).

As a result, the GRNs underlying nearly all processes of NC development have been addressed (M. Simões-Costa et al. 2015). NPB induction and NC establishment are the first key stages of NC development, during which certain genes become activated along the embryonic ectoderm. These genes exert positive feedback on each other to induce commitment to a certain fate, while also inhibiting factors functioning in adjacent domains, often via repressive feedback loops (Groves and LaBonne 2014; M. Simões-Costa et al. 2015). Subsequently, EMT and NC delamination, migration along defined pathways and fate segregation towards diverse derivatives are governed by separate networks (Baxter et al. 2010; Theveneau and Mayor 2012). Thus, dynamic transitions between complex GRNs take place in the developing NC.

Despite their overwhelming complexity and the continuous surge of additional data, GRNs underlying NC development are currently being represented diagrammatically and

lack any further processing. They are therefore very difficult to analyse, evaluate and refine. An increasing body of work on other systems has focused on mathematically modelling GRNs, as a way to better understand the molecular basis of complex developmental processes and to refine biologically derived networks. For instance, discrete mathematical modelling has been successfully used to investigate the GRNs governing HSC diversification towards different blood lineages (Krumsiek et al. 2011; Tian and Smith-Miles 2014). Furthermore, previously overlooked mechanisms underlying the patterning of the embryonic spinal cord have been derived using experimental approaches combined with systems of ODEs (Balaskas et al. 2012; Cohen et al. 2015). Finally, in the NC context, the zebrafish melanophore GRN was generated using iterative cycles of modelling and experimentation (Greenhill et al. 2011).

Overall, despite the wealth of available information regarding the molecular basis of NC development, the GRNs which are being constructed still lack rigorous analysis, which in other systems has been achieved via mathematical modelling.

6.1.2. Pigment cells as models to study GRNs

All vertebrate pigment cell lineages are derived from the NC. While mammals only possess melanocytes, in zebrafish three different types of pigment cells have been identified: black melanophores, light-reflecting iridophores and yellow xanthophores. Due to their inherent properties, these cells become excellent models to study the GRNs governing diversification of different lineages from the NC. Specifically, chromatophores are highly tractable, accessible and amenable to genetic manipulation. In addition, several reports identified the existence of common pigment cell precursors (Lopes et al. 2008; Curran et al. 2010; M. Nikaido, unpublished data), thus rendering these lineages appropriate to investigate the cellular and molecular dynamics of partial fate restriction processes, which underlie NC derivative specification according to the progressive fate restriction model.

Within the pigment cell field, but also in order to develop a holistic view of NC development, it is important to understand the molecular mechanisms guiding fate determination of the different pigment cell lineages. Although the GRN governing melanocyte development has been focused on, not only in zebrafish but also in mammals (Baxter et al. 2010; Greenhill et al. 2011; D'Mello et al. 2016), the molecular mechanisms underlying iridophore and xanthophore development remain largely elusive.

In this study, the aim was to employ a systems biology approach, taking advantage of both traditional experimental techniques and mathematical modelling, to generate the GRN governing iridophore development. Ongoing work in our research group, in

collaboration with Dr H. Schwetlick (University of Bath) and Dr A. Rocco (University of Surrey), focuses on merging the resulting GRN with the previously described melanophore GRN (Greenhill et al. 2011), to generate a much more complex network able to predict the process of fate segregation of a characterised bipotent progenitor, the melanoiridoblast (Curran et al. 2010). Furthermore, the methodology used here can be readily extended to mathematically model and, thus, refine the intricate GRNs guiding NC development.

6.2. The iridophore GRN

6.2.1. An iterative process of GRN refinement

In contrast to the genetic interactions guiding development of several NC derivative lineages, those underlying iridophore specification have not been identified. The widely supported in the NC field progressive fate restriction model dictates that derivatives become gradually specified from a series of common progenitors. Therefore, in order to improve our understanding of the molecular dynamics governing fate segregation, it is imperative to establish the interactions guiding specification towards individual fates. The lack of a GRN underlying iridophore fate choice thus impacts on our understanding of NC development. In this work, a series of experiments were conducted to establish a simple core model for iridophore specification (Fig. 6.1, model A). Moreover, iterations of mathematical modelling and experimental testing facilitated the refinement and expansion of this preliminary network to ultimately generate a more sophisticated and biologically accurate model D (Fig. 1.6).

Mathematical modelling has been incorporated in the experimental biology toolbox as a powerful method of converting assumptions into conclusions in a logical manner, not in an attempt to describe nature, but to more accurately describe our limited understanding of natural processes (Gunawardena 2014). Furthermore, this approach facilitates the derivation of formal assumptions, definition of expectations and devising new tests (Black 1989). Thus, mathematical modelling and experimentation cooperatively shed light on biological systems by asking relevant questions, testing assumptions and deriving testable hypotheses. An example of successful implementation of this integrated approach is previously published work in the group, which shed light on the GRN governing melanophore specification from the zebrafish NC and subsequent differentiation of this pigment cell lineage (Greenhill et al. 2011). This study demonstrates several advantages of using mathematical modelling as a tool to generate GRNs. It also advances the modelling methodology employed by Greenhill et al. by introducing physiologically relevant parameter sets in the mathematical description of the model, as well as by introducing a Monte Carlo algorithm that rigorously assesses the parameter

space to identify possible sensitivity of the modelling outputs to the pre-determined parameter sets.

Regarding predictions derived from the presented models, predicted inaccuracies in the initial models A, B and C were successfully eliminated. However, model D outputs still present with inconsistencies compared to experimental data. Specifically, in model D *mitfa* repression in the iridophore lineage could not be fully addressed, even though steps were taken to increase the accuracy of simulated *mitfa* expression dynamics compared to model C. Importantly, an uncharacterised factor R was incorporated in the adopted version of model D, which qualitatively produced appropriate predictions for *mitfa*. This addition was particularly instructive, as it established a set of requirements to be fulfilled by the TF that functions as factor R, for example when considering the expression pattern in the lineage of interest. It thus becomes possible to screen multiple alternative regulators and exclude those that do not meet the set requirements. Nevertheless, model D did not predict complete inhibition of *mitfa* expression in iridophores, thus further modifications of the model or re-evaluation of parameter constants are needed.

Another outstanding consideration is that, in models C and D, simulating loss of *ltk* function failed to predict any consequences in gene expression dynamics. Biologically, this was deemed inaccurate, as *ltk* mutant embryos present with gradual downregulation of iridogenic genes. Loss of *ltk* dependency in the model was a direct consequence of the *sox10/tfec* positive feedback loop effectively maintaining expression of all iridogenic genes even in the absence of Ltk signalling. Although preliminary experiments and theoretical testing presented here attempted to address whether *ltk* plays a role in the *sox10/tfec* positive feedback loop, results were inconclusive. Additional work is thus required to investigate this potential feature of the iridophore GRN, as well as to assess whether *sox10* and *tfec* do not in fact form a feedback loop. A subsequent model E should focus on resolving these two issues, with the aim to further increase the biological relevance of modelling predictions.

It is worth noting that the genetic interactions in model D closely resemble several of the typically encountered GRN structural motifs (refer to section 1.2) (Alon 2007). For instance, positive feedback loops are discerned, such as the ones between *tfec* and *ltk* as well as between *tfec* and *sox10*. Such positive feedback loops are vital for 'locking' iridophore fate choice through stabilising the expression of these major iridogenic genes. Additionally, model D displays various types of both coherent and incoherent feed forward loops (FFLs; Fig. 1.6). For example, *sox10* contributes to *pnp4a* activation through two distinct pathways involving different intermediate TFs, namely *tfec* and *mitfa* (Fig. 1.6; coherent FFL type 1). Moreover, *sox10* drives expression of both *tfec* and *mitfa*, while *tfec* mediates repression of *mitfa* (Fig. 1.6; incoherent FFL type 1). Because *tfec*

was also proposed to maintain *sox10* expression in this model, in addition to driving repression of the *sox10* target, *mitfa*, an incoherent type 3 FFL also arises. Finally, *tfec* input on *pnp4a* can be classified as an incoherent type 4 FFL, since *tfec* activates *pnp4a* as well as repressing its alternative activator, *mitfa*.

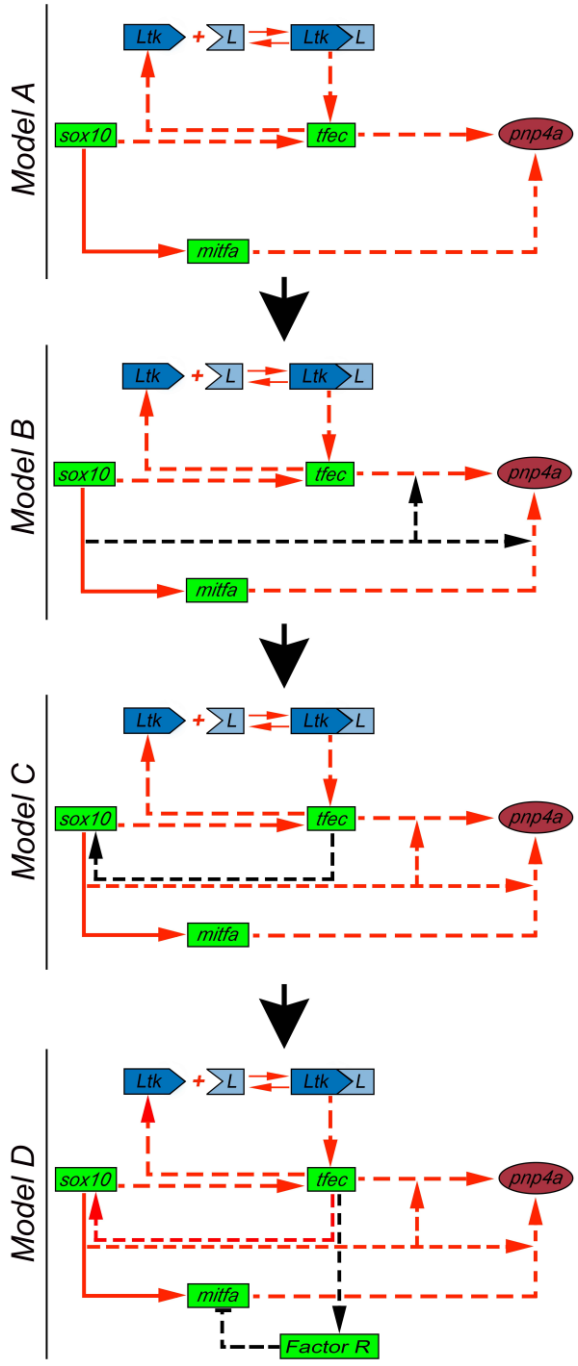


Figure 6.1. Iterative evolvement of the iridophore GRN. The progression from the experimentally derived model A, to model B, model C and, finally, model D. Red arrows indicate activatory interactions. Black arrows indicate the added interactions at each stage, including repression of *mitfa* by factor R in model D.

6.2.2. Novel roles for *sox10* in NC diversification

Using the aforementioned approach, several novel features of the iridophore GRN were described. In model A, *sox10* was highlighted as an essential feature of the GRN, responsible for maintaining expression of the presumed iridophore master regulator, *tfec*, during the transition from the premigratory NCC population to specified iridoblasts. In turn, *tfec* and the RTK-coding gene, *ltk*, were proposed to form a positive feedback loop crucial for survival of the lineage to differentiation stages. The differentiation gene, *pnp4a*, was shown to be downstream of *tfec* in the GRN. Surprisingly, *pnp4a* was also demonstrated to be a target of the melanophore master regulator, *mitfa*, exclusively during early stages of pigment cell specification, thus contradicting the previously established role for *pnp4a* as an early iridophore-specific marker (Curran et al. 2010).

This wiring of the GRN bears striking resemblances to that of both the melanophore and the sensory neuron GRNs (Fig. 6.2) (Greenhill et al. 2011; Prendergast and Raible 2014). Melanophore development relies upon *sox10* to induce *mitfa* expression, which subsequently activates melanophore differentiation genes, such as *dct*, *tyr* and *tyrp1*. Moreover, *mitfa* was shown to form a positive feedback loop with an unknown factor Y in order to maintain its own expression in developing melanophores, even after *sox10* has become downregulated (Greenhill et al. 2011). Intriguingly, factor Y has been indicated to correspond to the RTK-coding gene, *kit*, which is vital for *mitfa* maintenance and melanophore survival (Rawls & Johnson 2003; A. Lapedriza, PhD thesis).

Furthermore, similar regulatory mechanisms, studied in different model species, appear to be in place in the developing sensory neurons. Specifically, in zebrafish, *sox10* has been implicated in initiating expression of neurogenin, which codes for a bHLH TF necessary for neuronal development (Q. Ma, Kintner, and Anderson 1996; Carney et al. 2006). Several lines of evidence have suggested that the function of neurogenin family TFs is not sufficient to induce specification of sensory rather than sympathetic neurons, thus, they are not regarded as master regulators of the sensory neuron lineage (Pavan and Raible 2012). Instead, they are responsible for initiating the expression of sensory neuron-specific downstream TFs such as Brn3a and Islet1 (Pavan and Raible 2012; Prendergast and Raible 2014). Brn3a has been shown to regulate Trk RTK genes, vital for survival of different sensory neuron lineages (Huang et al. 1999; L. Ma et al. 2003; Lei et al. 2006). Although Trk feedback on neurogenin expression has not been explicitly reported in the sensory neuron context, it was recently shown that human TRKB regulates neurogenin expression in the pancreas (Shamblott et al. 2016), suggesting that such an interaction might be in place for neuronal survival.

Overall, a pattern appears to be established, in which *sox10* triggers bHLH TFs essential for specification of different NC derivatives. The expression of these TFs is shown or

presumed to be maintained via establishment of positive feedback with the appropriate RTK genes for each lineage. This conclusion is reached based on previous data for the melanophore and neuronal lineages and is reinforced by the identified interactions within the iridophore GRN.

Although these networks present with several similarities, one important difference is that *sox10* expression is maintained throughout iridophore development (this work), while both in melanophores and in sensory neurons undergoing specification the gene plays only a transient role and becomes downregulated (Cheng et al. 2000; Greenhill et al. 2011). *sox10* repression in the developing melanophore lineage has been suggested to occur via *hdac1*-mediated repression of the *mitfa* input on *sox10*. Although the mechanism by which *sox10* is inhibited in the sensory neuron lineage has not been determined, the observed analogies between the networks might suggest that a similar set of interactions is in place. Furthermore, contrary to what has been shown for melanophore differentiation genes, on which *sox10* exerts a negative influence (Greenhill et al. 2011), expression of the iridophore gene *pnp4a* was found to require *sox10* function (this work). Therefore, a novel role for *sox10* in promoting specification and differentiation of an NC-derived lineage is established. Intriguingly, this role resembles the gene's previously investigated function in glia development (Britsch et al. 2001).

An additional feature of *sox10* regulation proposed over the course of this project concerned the existence of a positive feedback loop between *tfec* and *sox10*. This was considered important for iridophore-specific maintenance of *sox10* and was supported by preliminary data from overexpression studies. Interestingly, it has been previously indicated that *mitfa* is similarly able to activate *sox10* expression in melanophores (Greenhill et al. 2011). The newly developed *tfec* mutant lines will prove valuable tools for definitively evaluating the presence of the *sox10/tfec* feedback loop.

Conclusively, the major features of the iridophore GRN closely resemble those of GRNs guiding specification of other NC derivatives. As has been previously shown, *sox10* is important for initiation of the regulatory cascade but, importantly, remains involved in iridophore specification and likely also differentiation. It should not be disregarded that GRNs present with common motifs (Alon 2007; Burda et al. 2011), which could lead to analogies in structural features. However, the striking similarities between members of the outlined GRNs and their respective interactions indicate that the described networks evolved closely and that their parallels could potentially be harnessed to aid the identification of novel features for each derivative.

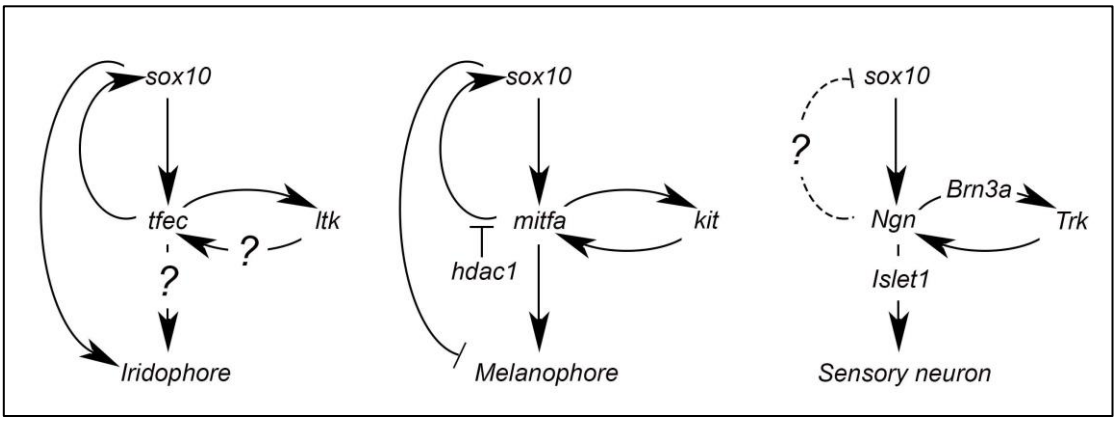


Figure 6.2. Direct comparison of the iridophore, melanophore and sensory neuron specification GRNs. The core features of the iridophore GRN (this work) are compared with those of the melanophore and the sensory neuron GRNs. Data presented here indicated that additional cofactors (represented by question marks) are required for *ltk* to upregulate *tfec* and for *tfec* to activate iridophore differentiation genes, in analogy with mammalian neurogenins (Ngn) acting through Brn3a and Islet1 to induce specification of the sensory neuron lineage. Judging by the similarities of these networks, it is plausible that repression of *sox10* in developing sensory neurons may be mediated by Ngn and, perhaps, other factors, although this feature remains to be proven.

6.2.3. The role of *foxd3* during pigment cell development

Over the course of this project, a previously uncharacterised *foxd3* mutant line, *foxd3*^{sa20726}, was analysed with the view to settle the pre-existing controversy concerning the gene's role in pigment cell development. As stated, Curran et al. developed a model, according to which *foxd3*-dependent *mitfa* repression aids iridophore specification, at the expense of melanophores (Curran et al. 2010). Nevertheless, data presented here and previously published work are not consistent with this model, as there is no increase in the numbers of *mitfa* positive cells or differentiated melanophores in *foxd3* mutants (Lister et al. 2006; Montero-Balaguer et al. 2006; Stewart et al. 2006; this work), although this conclusion is disputed by one study (Hochgreb-Hägele and Bronner 2013). Loss of *foxd3* function does, however, consistently result in a partial reduction of iridophores manifesting solely along the dorsal stripe. Thus, it was considered interesting to further investigate the role of *foxd3* in pigment cell development, focusing on the melanophore and the iridophore lineages.

To that effect, the expression patterns of both *tfec* and *mitfa* were analysed in *foxd3* mutants over a time course. Interestingly, results indicated a defect in specification of both melanoblasts and iridoblasts, which potentially related to the timing of the process. Specifically, as was reported in previous studies (Montero-Balaguer et al. 2006; Stewart et al. 2006), migration of *mitfa* positive cells was delayed upon loss of *foxd3* function. Curiously, however, a significant number of melanoblasts expressing *mitfa* were absent from the medial migratory pathway, while cells could be observed along the lateral pathway, which is normally occupied approximately 3 hours after the medial. This previously unacknowledged feature suggested that a distinct subset of melanoblasts, likely of different origin, were either specified with a significant delay, or never became specified.

Moreover, in *foxd3* mutants very few *tfec* positive cells entered the migratory pathways. Instead, prolonged expression along the dorsal trunk was observed. This resembled the *tfec* and *ltk* expression patterns in *sox10* mutants (Lopes et al. 2008; this work), except the number of "trapped" progenitors appeared to be smaller in *foxd3* mutants. Since the *tfec* and *ltk* transcripts have been suggested to label the partially fate restricted tetrapotent progenitor, chromatoglioblast (Lopes et al. 2008; this work), it was postulated that the *tfec* positive population observed in *foxd3* mutant embryos likely consisted of a less abundant, partially restricted cell type. Together, these data suggested that the *mitfa* positive cells missing from the migratory pathways were either temporarily or permanently restricted along the premigratory NC domain and most likely correspond to at least a subset of *tfec* expressing cells identified in this region. In the future, co-expression studies should confirm this overlap, but it seems likely that melanoiridoblasts,

positive for both *mitfa* and *tfec*, remain trapped in the premigratory domain in *foxd3* mutants.

The roles of *foxd3* in pigment cell development require further investigation to clarify whether the restriction of progenitors along the premigratory domain is only temporary. It is considered likely that the missing melanoblasts did begin to migrate later in development, as by 30 hpf the phenotype is almost completely absent. However, it is likely that a subpopulation of iridoblasts never recovered from a potential delay and either underwent apoptosis, or switched fate. Although a fate switch towards the melanophore lineage is unlikely, due to reasons discussed above, it cannot be excluded that it occurs transiently in a small subset of common progenitors, and therefore it is difficult to detect a corresponding increase in melanophore numbers. To address this, the precise number of migrating *mitfa* positive melanoblasts should be scored in different sections of the trunk of WT and *foxd3* mutant embryos. It may also prove useful to conduct apoptosis assays in order to establish whether trapped cells biased towards an iridophore fate are lost by apoptosis.

In different model organisms, Foxd3 has been shown to assume distinct, vital roles in various stages of NC development, including NC induction, survival of multipotent progenitors, delamination of NCCs and, ultimately, fate choice (Kos et al. 2001; Sasai, Mizuseki, and Sasai 2001; Cheung et al. 2005; Montero-Balaguer et al. 2006; Teng et al. 2008; Mundell and Labosky 2011; M. B. Weiss et al. 2014). Data presented here support the previously established delay in melanoblast specification upon loss of *foxd3* function, as well as the defect in iridophore development, which potentially results from prolonged delay in specification of the lineage. Moreover, a body of work suggests that Foxd3 is able to repress Mitfa *in vivo* and *in vitro* and that, in zebrafish, this interaction aids iridoblast specification (Ignatius et al. 2008; Thomas and Erickson 2009; Curran, Raible, and Lister 2009; Curran et al. 2010). Finally, in this work, it was shown that *mitfa* repression should be sustained during iridophore differentiation stages, however *foxd3* was indicated to not fulfil key requirements for the repressor of *mitfa* in this context.

Overall, it becomes apparent that *foxd3* exerts distinct transcriptional effects during different developmental stages but also according to the cell type in which it is expressed (H.-C. Lee et al. 2006; Y. Liu and Labosky 2008). These spatiotemporally distinct effects might be explained by considering *foxd3* as a potential pioneer factor, functioning to modify chromatin and trigger the deposition of epigenetic marks by recruiting appropriate complexes, so that genes remain transcriptionally inactive but become poised to respond to stimuli from the microenvironment (Zaret and Carroll 2011; Krishnakumar et al. 2016).

6.2.4. *tfec* in pigment cell development

The fact that *tfec* encodes a TF belonging to the same family as Mitfa, is expressed throughout iridophore development (Lister et al. 2011; this work) and its function is indispensible in the lineage (J. Lister, unpublished data; this work) aided the conclusion that *tfec* is likely the iridophore master regulator. Curiously, however, ectopically expressing WT *tfec* has thus far not succeeded in rescuing the iridophore phenotype manifesting in *tfec* mutant embryos (J. Lister, unpublished data). On the contrary, ectopic expression of *mitfa* rescues melanophore development in *mitfa* mutants (Elworthy et al. 2003). Although there are several technical explanations as to why *tfec* expression did not appear to be sufficient for iridophore development, these results indicate that *tfec* may not fulfil the requirements of a master regulatory gene. Instead, it seems possible that the function of *tfec* is analogous to that of the pro-neural TF, neurogenin.

In particular, widespread expression of *tfec* across the premigratory NC domain at 18 hpf (Lister et al. 2011; this work) indicates that it might be functioning in other cell lineages and that its expression potentially only becomes restricted to the iridophore lineage upon activation of additional factors. This behaviour is in line with neurogenin expression and function in common progenitors of PNS components, including precursors of autonomic neurons and glial cells (Lo et al. 2002; Zirlinger et al. 2002), as well as neurogenin-dependent upregulation of sensory neuron-specific TFs such as Brn3a and Islet1 (Fig. 6.2) (Pavan and Raible 2012; Prendergast and Raible 2014). Overall, it is clear that several key players of the iridophore GRN are still missing. To better understand development of the lineage and how the iridophore GRN relates to that governing neuronal specification, it is necessary to identify additional genes downstream of *tfec* and to investigate how they might function to bias progenitors towards the iridophore fate.

Interestingly, this work indicated a previously unrecognised interaction between *mitfa* and *tfec*, namely a repressive feedback loop that likely biases the melanoiridoblast towards either a melanophore or an iridophore fate. Specifically, it was shown that loss of *mitfa* function resulted in a significant increase in the number of cells expressing *tfec* from early stages of specification, suggesting that downregulation of *tfec* from premigratory NCCs in the WT is *mitfa*-dependent. Moreover, preliminary results using *tfec* mutants have indicated a delay in the migration of *mitfa* positive NC derivatives. Intriguingly, at pigment cell differentiation stages the melanophore population appeared to increase in *tfec* mutants compared to WT siblings, a phenotype that has yet to be quantified. Together, these data suggested that *mitfa* represses *tfec* expression in developing NCCs, while *tfec* initially functions to induce timely specification of melanoblasts and subsequently to bias a subset of progenitors towards an iridophore

rather than a melanophore fate. Potential heterodimerisation between Tfec and Mitfa as well as the dynamics and the functional implications of this interaction during NC development (refer to section 4.7) remain to be determined. Moreover, further experimentation using *tfec* mutants is required to assess definitively the potential implications of loss of *tfec* function in relation to melanoridoblast fate choice.

These results offer an alternative to the currently accepted model for segregation of the melanophore and the iridophore lineages from the melanoiridoblast, under which *foxd3* represses *mitfa* to aid iridophore development (Curran, Raible, and Lister 2009; Curran et al. 2010). In agreement with published literature, results argued against this model, on the basis of the number of mature melanophores not increasing in *foxd3* mutants (Lister et al. 2006; Montero-Balaguer et al. 2006; Stewart et al. 2006; this work). Data presented here propose that mutual repression between *tfec* and *mitfa* is likely responsible for biasing the melanoiridoblast towards either the iridophore or the melanophore fate.

6.3. Systems biology as a new tool to study NC GRNs

6.3.1. Requirement for mathematical modelling

The vast complexity of gene regulatory interactions governing the various stages of NC development has been identified in the literature (M. Simões-Costa et al. 2015). Due to the profound intricacy and the dynamic nature of these interactions, analysing, refining and evolving the GRNs is not possible without the appropriate tools. Mathematical modelling has proven useful in the context of pigment cell development (Greenhill et al. 2011; this work), as well as in non-NC contexts (Krumsiek et al. 2011; Balaskas et al. 2012; Webb et al. 2016), as it allows for using the existing assumptions to derive predictions and to formulate testable hypotheses.

As can be observed from the NC specification network (section 1.3.2), in several cases the activatory inputs on several genes occur redundantly. For instance, Snai2 is activated by Snai1 and by Sox10, while Snai1 has also been reported to upregulate the expression of Sox10 (Aybar, Nieto, and Mayor 2003; Honoré, Aybar, and Mayor 2003). Since Snai1 does not seem to be Sox10-dependent, mathematically modelling this GRN would predict that in the absence of Sox10, Snai1 would still be able to induce the expression of Snai2. However, experimental data argue that expression of Snai2 is repressed in the absence of Sox10 (Honoré, Aybar, and Mayor 2003), immediately rendering this aspect of the GRN inaccurate. Thus, the hypothesis that Snai1 only acts through Sox10 to upregulate Snai2 can be formulated, and subsequently tested, for example by overexpressing Snai1 in Sox10 mutants to assess whether Snai2 becomes ectopically expressed.

The Snai1/Snai2/Sox10 was a simple and relatively obvious example, useful to illustrate how compiling data from different sources while lacking methods of validating their accuracy can be dangerously misleading. Using the established approach which was undertaken in this study, as well as the extensive body of experimental observations stemming from decades of research on the molecular mechanisms guiding NC development, the underlying GRNs have the potential to evolve even further.

6.3.2. Advanced experimental approaches to generate accurate assumptions

To generate reliable predictions using mathematical modelling, it is crucial that adequate experimental data lead to valid interpretations with regard to network structure (Gunawardena 2014). Research on the NC has benefited from sophisticated experimental approaches, which have greatly advanced our understanding of molecular mechanisms. These approaches, carried out using diverse model organisms, have led to identifying novel factors involved in processes of interest, determining direct interactions and obtaining quantitative data. Moreover, they have allowed for examining gene expression dynamics *in vivo*. In this section, some of the experimental techniques employed in the NC field, which are proving valuable when constructing developmental GRNs, are discussed.

One important consideration when investigating GRNs, is to have a comprehensive view of the factors involved in a particular process. As discussed in previous sections, our understanding of key members of the iridophore GRN appears to be incomplete. Specifically, molecules acting downstream of *tfec* to induce iridophore specification, similar to Brn3a and Islet1 acting downstream of neurogenins in sensory neuron specification, remain unidentified. Furthermore, overexpression studies presented here highlighted that important co-factors required for *pnp4a* activation and for *ltk*-dependent upregulation of iridogenic genes have yet to be identified.

In previous years, genome-wide mutagenesis screens have proven invaluable to identify novel genes involved in a process of interest (R N Kelsh et al. 1996; Odenthal, Haffter, et al. 1996; Amsterdam et al. 1999). However, this method can be very time-consuming, and even impossible in cases where mutations result in embryonic lethality. Microarray studies have also been used to identify global changes in gene expression between WT and embryos with loss of gene function, in which a phenotype of interest manifests (A. Y. H. Leung et al. 2005). More recently, RNA-Seq has gained support as it allows for high-throughput transcript sequencing, presents with significantly increased resolution and lower background noise compared to microarray assays as well as with the capacity to distinguish between different isoforms of a gene. In addition, as opposed to previous methods, this methodology requires a relatively low amount of transcript (Z. Wang,

Gerstein, and Snyder 2009). RNA-Seq has recently been applied in the cNC context, where a transgenic line allowed for isolation and transcriptome profiling of a specified population of migrating cells (Simoes-Costa et al. 2014). In the future, the generation of additional transgenic lines, selectively labelling subpopulations of the NC or of NC-derived lineages will allow for this technique to become widely adopted.

In order to study the functions of identified genes of interest, as well as their potential interactions with other members of the network, loss of function and overexpression experiments have been primarily used (Pohl & Knöchel 2001; Honoré et al. 2003; Sato et al. 2005; Plouhinec et al. 2014; this work). As outlined in previous sections, overexpression experiments can be difficult to interpret, as the context is not physiologically relevant. Therefore, it is plausible that false positives and negatives are often generated. Inducing of loss of function thus represents a more reliable approach to identify gene functions and intergenic interactions, although may still prove misleading where gene expression is regulated redundantly by different factors.

To that effect, morpholino technology has been frequently used over several years, as it offered a fast way to screen for potential phenotypes (Nasevicius and Ekker 2000; Nagatomo and Hashimoto 2007; M. S. Simões-Costa et al. 2012). However, an increasing amount of evidence has drawn attention to the fact that non-specific effects can occur using this approach (Robu et al. 2007; Eisen & Smith 2008; Schulte-Merker & Stainier 2014). Therefore, methods for achieving targeted mutagenesis have been widely adopted, with CRISPR/Cas9 site-directed mutagenesis recently receiving increasing support (Gaj et al. 2013; Hwang et al. 2013).

As was discussed in this work, CRISPR/Cas9 can be efficiently used to mutagenise a specified genomic region with minimal, if any, non-specific effects. This type of genomic manipulation relies upon NHEJ repair mechanisms. Moreover, the CRISPR/Cas9 system, in addition to TALEN nucleases, can be efficiently used to induce desired alterations, for instance with the aim of modelling a known human disease, as well as to insert several kilobase-long sequences into the genome (Kimura et al. 2015; Hoshijima, Jurynec, and Grunwald 2016). The latter application is especially applicable for GRN research, as it allows for generation of transgenic lines in which fluorescent proteins are induced in a lineage-specific manner, or are under the control of specified regulatory regions. Such tools render lineage tracing studies and *in vivo* characterisation of cellular processes using time-lapse microscopy feasible (Rodrigues et al. 2012; M. S. Simões-Costa et al. 2012).

Upon successfully ablating the function of a gene of interest, downstream effects on potential targets can be examined using different techniques. Conventional *in situ* hybridisation has been used widely in all species as a way of assessing spatio-temporal

gene expression changes in the loss of gene function context, compared with WT embryos, but also following overexpression. Recently, advanced methods for detecting transcript presence, as well as co-expression of two to four genes in whole-mount embryos have been developed. Specifically, fluorescent hybridisation chain reaction (McLennan et al. 2015), as well as RNAscope (Gross-Thebing et al. 2014; this work) have been employed. Importantly, both of these techniques allow for quantitation of gene expression levels based on fluorescence emission, thus addressing the previously discussed requirements of being able to distinguish cells of interest, while also measuring expression levels within these cells (refer to section 3.4.4). Another well-established method for quantitatively assessing gene expression is q-RT PCT, although this technique involves dissociation of samples and, therefore, does not allow for differentiating between cell types.

Having assessed intergenic relationships, specifically whether ablation of the function of a gene leads to repression or ectopic activation of a downstream factor, it remains to be determined whether the identified interaction occurs directly or indirectly. In other words, it is imperative to know whether a TF binds to the regulatory elements of a gene, or whether its effects are elicited through intermediate factors. The nature of identified interactions was not addressed in the context of the iridophore GRN due to unavailability of required resources at the time, although future studies should further investigate this matter.

ChIP-Seq is the most widely used approach for identification of direct protein-DNA interactions (Furey 2012), and has been applied to study the precise molecular mechanisms underlying aspects of NC development (B. Li et al. 2009; M. S. Simões-Costa et al. 2012; Rada-Iglesias et al. 2012). ChIP-Seq experiments, although very informative, require careful design and analysis to ensure that experimental artefacts are eliminated (Kidder, Hu, and Zhao 2011). One important consideration is the existence of sensitive antibodies, which in zebrafish remain largely unavailable. In the absence of antibodies specific to a protein of interest, it is possible to overexpress a tagged protein, subsequently detectable using a tag-specific antibody (Kolodziej et al. 2009). Ongoing work in our research group focuses on generation of transgenic zebrafish lines using the CRISPR/Cas9 system, in which endogenous TFs of interest will be tagged (Dr T. Subkhankulova, unpublished data). Therefore, likely physiologically irrelevant effects resulting from overexpressing tagged proteins are avoided.

It should furthermore be considered how experimental conditions might affect ChIP-Seq results. In two recent publications, Foxd3 binding was demonstrated on distinct sets of target genes possibly due to differing culture conditions (Respuela et al. 2016; Krishnakumar et al. 2016; Sweet et al. 2016). Finally, it is occasionally difficult to interpret

ChIP-Seq data, as it is not trivial to assign enhancers found to interact with the TF of interest to the appropriate regulated gene. Overall, ChIP-Seq experiments remain challenging but continuous advances render them progressively more appropriate to identify direct interactions *in vivo*, therefore aiding the generation of GRNs incorporating fewer assumptions. An additional methodology used to determine the ability of a TF to bind a specific DNA region of interest is the electrophoresis mobility shift assay (EMSA), although this experiment is performed *in vitro* and thus not in a physiologically relevant context (Thomas and Erickson 2009; Plouhinec et al. 2014).

The outlined approaches have contributed a wealth of information with regard to molecular mechanisms, valuable for constructing GRNs. Mathematical modelling is bound to be useful for effectively compiling the existing data and further evolving the resulting GRNs.

6.4. The melanoiridoblast GRN

The iridophore GRN generated in this study made an important contribution to the development of the melanoiridoblast network, which is currently investigated in our group, in collaboration with Dr H. Schwetlick (University of Bath) and Dr A. Rocco's group (University of Surrey). The objective is to undertake the systems biology approach described here, in order to refine and expand the GRN to better understand the process of fate segregation of this previously characterised bipotent progenitor (Curran et al. 2010).

To build the preliminary version of this complex GRN (Fig. 6.3), the melanophore (Greenhill et al. 2011) and the iridophore (this work) GRNs were combined, and the literature was searched for known interactions between members of these core networks with other factors active during specification stages. Additional data generated by previous members of the group (Dr A. Lapedriza, Dr L. Vibert), as well as deriving from overexpression studies (Dr T. Subkhankulova) were incorporated. The resulting GRN is being mathematically modelled using a system of ODEs (Dr G. Aquino, University of Surrey), initiating the iterative process of modelling and experimentation which has been used to generate the individual GRNs governing melanophore and iridophore development (Greenhill et al. 2011; this work).

As has been pointed out in the literature (Gunawardena 2014; Kirk, Babtie, and Stumpf 2015), mathematical modelling cannot be considered sufficient to produce complete GRNs, which accurately reflect known biology. To improve the use of this tool, in particular when using ODEs, reliable sets of quantitative data over a time course are required. Importantly, predictions from modelling the melanoiridoblast GRN will be tested using quantitative single cell data (Dr T. Subkhankulova, unpublished data).

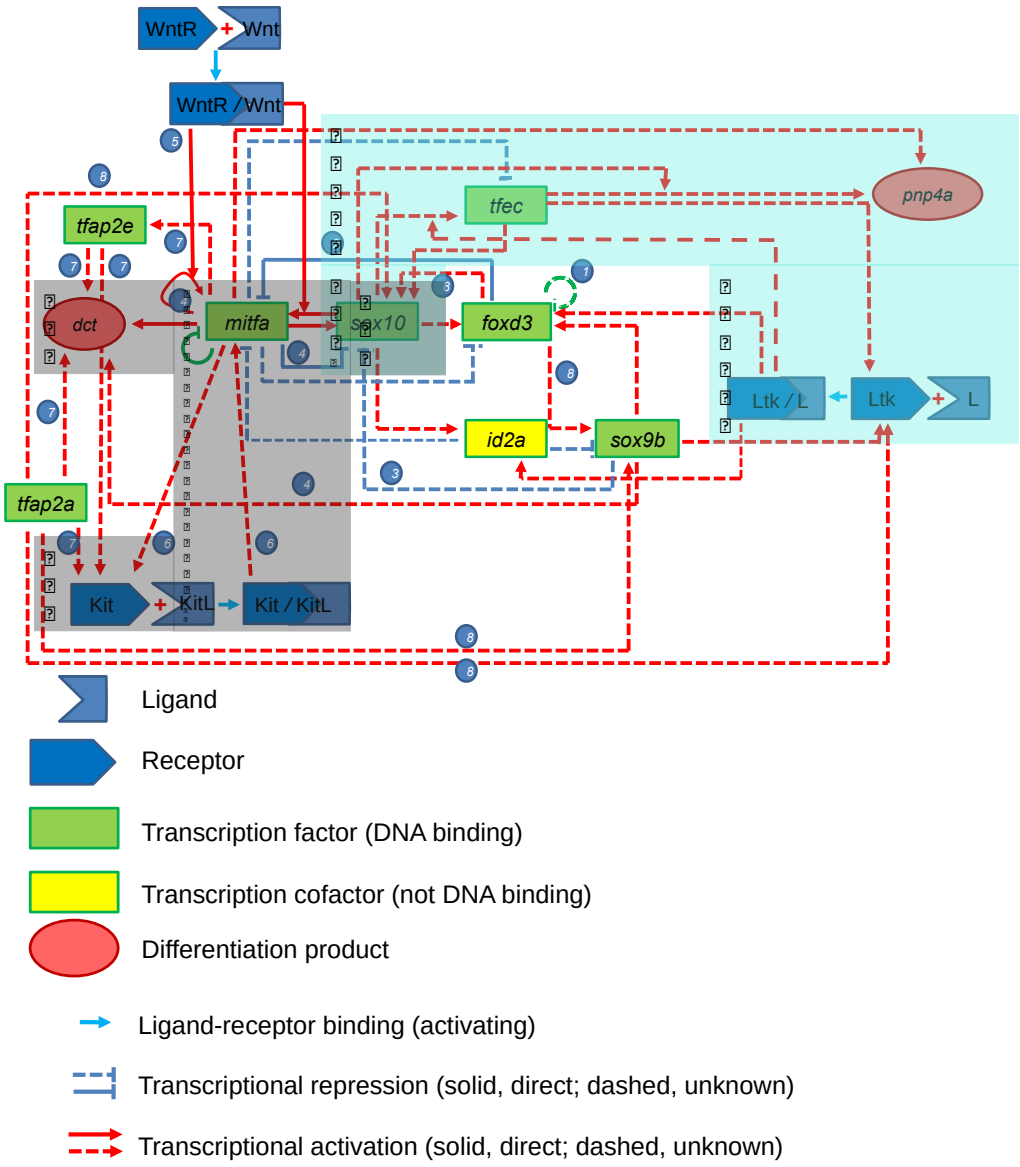


Figure 6.3. The melanoiridoblast preliminary GRN. The core model describing fate switch of the melanoiridoblast, based on data from published literature, previous work by members of our group, the iridophore GRN presented here and overexpression studies by Dr T. Subkhankulova. The highlighted areas correspond to the key players of the melanophore GRN (black; Greenhill et al. 2011) and the iridophore GRN (blue; this work). References: **1**, Pohl & Knöchel 2001; **2**, Curran et al. 2009; **3**, Shakhova et al. 2015; **4**, Greenhill et al. 2011; **5**, Dorsky et al. 2000; L. Vibert, PhD thesis; **6**, A. Lapedriza, PhD thesis; **7**, O'Brien et al. 2004; Knight et al. 2004; Van Otterloo et al. 2012; **8**, Arduini et al. 2009.

Specifically, NanoString technology (Jung and Sohn 2014) is being used to quantify absolute expression levels of genes of interest in NC-derived single cells from *sox10*:GFP embryos, isolated using FACS at presumed key stages to capture development of common pigment cell precursors and of individual chromatophore lineages. Obtaining this type of data should also enable the use of network inference algorithms, thus addressing another concern arising with the use of mathematical modelling, namely the uncertainty with regard to the validity of adopted network structures (Kirk, Babtie, and Stumpf 2015).

Overall, this project takes advantage of single-cell technology to capture quantitatively the gene expression dynamics taking place as NCC progenitors become progressively fate restricted. The gene expression signatures obtained are expected to aid identification of the bipotent melanoiridoblast and of fate restricted iridoblasts and melanoblasts over a developmental time course, thus providing the required experimental input for iteratively evolving the preliminary GRN.

6.5. Summary

The vertebrate NC is an excellent model for studying the GRNs guiding the sequential stages of development of this highly dynamic cell population, including diversification of NC derivatives. The three chromatophore lineages of zebrafish pose an attractive system to study the molecular mechanisms underlying fate segregation from the NC. Although the GRN guiding melanophore development has been previously investigated using an iterative systems biology approach (Greenhill et al. 2011), the iridophore and xanthophore networks remain elusive, thus hindering our understanding of chromatophore fate segregation. This project aimed to characterise the GRN guiding iridophore development, using the mathematical modelling and experimentation approach employed for the melanophore model.

Results described here first identified for the first time the core intergenic interactions underlying iridoblast specification and, subsequently, revealed novel regulatory features which would not have been predicted by biology alone. For instance, mathematical modelling predicted that *sox10* expression should be maintained in the iridophore lineage for the differentiation marker *pnp4a* to not become downregulated. Indeed, subsequent experimentation confirmed that *sox10* expression was maintained throughout iridophore development, leading to the proposal of a positive feedback loop between *tfec* and *sox10*, interestingly resembling the feedback loop between *mitfa* and *sox10*, previously characterised in melanophore development (Greenhill et al. 2011). Furthermore, *sox10*-dependent upregulation of *pnp4a* contrasts with the role of *sox10* in the melanophore lineage, where the TF was found to repress differentiation genes.

Furthermore, the role of *foxd3* as the *mitfa* repressor in the iridophore lineage was investigated. Although this result is at odds with the currently accepted model in the field (Curran et al. 2010), results presented here indicated that *foxd3* is insufficient to repress *mitfa* in the iridophore lineage. Further investigations suggested that *foxd3* may be important for the specification of an NC-derived fate-restricted progenitor of both melanophores and iridophores. Moreover, the developmental defects observed upon loss of *foxd3* function likely owed themselves to a delay in the specification of this progenitor, which only temporarily impacted on melanophores but permanently affected the iridophore lineage. Whether iridophore precursors switched to the melanophore fate due to de-repression of *mitfa* over a narrow time window, or whether they were lost by apoptosis remains to be determined. In addition, future work should further investigate the potential role of *tfec* as the inhibitor of *mitfa*, at least in the context of the melanoiridoblast, but likely also in mature iridophores.

Over the course of this project, the iterative systems biology approach was further developed and additional experimental and computational tools were established. While further work is required to determine direct interactions and to quantify gene expression dynamics, as well as to eliminate persisting inaccuracies in the final model D, this GRN is considered a more accurate representation of the interactions occurring in the iridophore lineage, when compared to an entirely experimentally determined model A. Importantly, the iridophore GRN has been used to construct the network guiding melanoiridoblast fate choice. As the presented approach appears to readily identify inaccuracies in the experimentally-derived GRNs, thus exposing assumptions and aiding the formulation of improved hypotheses, it would be interesting to apply it to the highly complex and dynamic networks governing NC development, which still lack adequate processing and methods of evaluation.

Bibliography

Abrams, J M, K White, L I Fessler, and H Steller. 1993. “Programmed Cell Death during Drosophila Embryogenesis.” *Development (Cambridge, England)* 117 (1): 29–43.

Adam, Rene C, Hanseul Yang, Shira Rockowitz, Samantha B Larsen, Maria Nikolova, Daniel S Oristian, Lisa Polak, et al. 2015. “Pioneer Factors Govern Super-Enhancer Dynamics in Stem Cell Plasticity and Lineage Choice.” *Nature* 521 (7552): 366–70.

Adameyko, Igor, and Francois Lallemend. 2010. “Glial versus Melanocyte Cell Fate Choice: Schwann Cell Precursors as a Cellular Origin of Melanocytes.” *Cellular and Molecular Life Sciences*.

Adameyko, Igor, Francois Lallemend, Jorge B Aquino, Jorge A Pereira, Piotr Topilko, Thomas Müller, Nicolas Fritz, et al. 2009. “Schwann Cell Precursors from Nerve Innervation Are a Cellular Origin of Melanocytes in Skin.” *Cell* 139 (2). Elsevier Ltd: 366–79.

Agarwal, R. P., and D. O’Regan. 2009. *Ordinary and Partial Differential Equations*. New York, NY: Springer New York.

Alon, Uri. 2007. “Network Motifs: Theory and Experimental Approaches.” *Nature Reviews Genetics* 8 (6). Nature Publishing Group: 450–61.

Amsterdam, A, S Burgess, G Golling, W Chen, Z Sun, K Townsend, S Farrington, M Haldi, and N Hopkins. 1999. “A Large-Scale Insertional Mutagenesis Screen in Zebrafish.” *Genes & Development* 13 (20): 2713–24.

An, Min, Rushu Luo, and Paul D Henion. 2002. “Differentiation and Maturation of Zebrafish Dorsal Root and Sympathetic Ganglion Neurons.” *The Journal of Comparative Neurology* 446 (3): 267–75.

Ando, Koji, Shigetomo Fukuhara, Nanae Izumi, Hiroyuki Nakajima, Hajime Fukui, Robert N. Kelsh, and Naoki Mochizuki. 2016. “Clarification of Mural Cell Coverage of Vascular Endothelial Cells by Live Imaging of Zebrafish.” *Development* 143 (8).

Andouche, Aude, Yann Bassaglia, Sébastien Baratte, and Laure Bonnaud. 2013. “Reflectin Genes and Development of Iridophore Patterns in Sepia Officinalis Embryos (Mollusca, Cephalopoda).” *Developmental Dynamics : An Official Publication of the American Association of Anatomists* 242 (5): 560–71.

Arduini, Brigitte L, Kevin M Bosse, and Paul D Henion. 2009. “Genetic Ablation of Neural Crest Cell Diversification.” *Development (Cambridge, England)* 136 (12): 1987–94.

Aybar, Manuel J, M Angela Nieto, and Roberto Mayor. 2003. “Snail Precedes Slug in the Genetic Cascade Required for the Specification and Migration of the Xenopus Neural Crest.” *Development (Cambridge, England)* 130 (3). The Company of Biologists Ltd: 483–94.

Bagnara, J T, J Matsumoto, W Ferris, S K Frost, W A Turner, T T Tchen, and J D Taylor. 1979. “Common Origin of Pigment Cells.” *Science* 203 (4379): 410–15.

Bagnara, Joseph T, Philip J Fernandez, and Royozo Fujii. 2007. “On the Blue Coloration of Vertebrates.” *Pigment Cell & Melanoma Research* 20 (1): 14–26.

Balaskas, Nikolaos, Ana Ribeiro, Jasmina Panovska, Eric Dessaud, Noriaki Sasai, Karen M Page, James Briscoe, and Vanessa Ribes. 2012. “Gene Regulatory Logic for Reading the Sonic Hedgehog Signaling Gradient in the Vertebrate Neural Tube.” *Cell* 148 (1–2): 273–84.

Ballerstein, K., U.U. Haus, J.A. Lindquist, T. Beyer, B. Schraven, and R. Weismantel. 2013. “Discrete, Qualitative Models of Interaction Networks.” *Front Biosci* 1 (5): 149–66.

Barembaum, Meyer, and Marianne E. Bronner. 2013. “Identification and Dissection of a Key Enhancer Mediating Cranial Neural Crest Specific Expression of Transcription Factor, Ets-1.” *Developmental Biology* 382 (2): 567–75.

Baroffio, A., E. Dupin, and N M Le Douarin. 1991. “Common Precursors for Neural and Mesectodermal Derivatives in the Cephalic Neural Crest.” *Development (Cambridge, England)* 112 (1): 301–5.

Baroffio, A, E Dupin, and N M Le Douarin. 1988. “Clone-Forming Ability and Differentiation Potential of Migratory Neural Crest Cells.” *Proceedings of the National Academy of Sciences of the United States of America* 85 (14): 5325–29.
Baxter, Laura L, R Travis Moreland, Anh-Dao Nguyen, Tyra G Wolfsberg, and William J Pavan. 2010. “A Curated Online Resource for SOX10 and Pigment Cell Molecular Genetic Pathways.” *Database : The Journal of Biological Databases and Curation* 2010 (January): baq025.
Baynash, A G, K Hosoda, A Giaid, J A Richardson, N Emoto, R E Hammer, and M Yanagisawa. 1994. “Interaction of Endothelin-3 with Endothelin-B Receptor Is Essential for Development of Epidermal Melanocytes and Enteric Neurons.” *Cell* 79 (7): 1277–85.
Belle, A., A. Tanay, L. Bitincka, R. Shamir, and E. K. O’Shea. 2006. “Quantification of Protein Half-Lives in the Budding Yeast Proteome.” *Proceedings of the National Academy of Sciences* 103 (35): 13004–9.
Bellmeyer, Amy, Jessica Krase, Julie Lindgren, and Carole LaBonne. 2003. “The Protooncogene c-Myc Is an Essential Regulator of Neural Crest Formation in Xenopus.” *Developmental Cell* 4 (6): 827–39.
Ben-Neriah, Y, and A R Bauskin. 1988. “Leukocytes Express a Novel Gene Encoding a Putative Transmembrane Protein-Kinase Devoid of an Extracellular Domain.” *Nature* 333 (6174): 672–76.
Bennett, Dorothy C., and M. Lynn Lamoreux. 2003. “The Color Loci of Mice - A Genetic Century.” *Pigment Cell & Melanoma Research* 16 (4): 333–44.
Bernards, A, and S M de la Monte. 1990. “The Ltk Receptor Tyrosine Kinase Is Expressed in Pre-B Lymphocytes and Cerebral Neurons and Uses a Non-AUG Translational Initiator.” *The EMBO Journal* 9 (7): 2279–87.
Betancur, P., M. Bronner-Fraser, and T. Sauka-Spengler. 2010. “Genomic Code for Sox10 Activation Reveals a Key Regulatory Enhancer for Cranial Neural Crest.” *Proceedings of the National Academy of Sciences of the United States of America* 107 (8): 3570–75.
Betancur, Paola, Marianne Bronner-Fraser, and Tatjana Sauka-Spengler. 2010. “Assembling Neural Crest Regulatory Circuits into a Gene Regulatory Network.” *Annual Review of Cell and Developmental Biology* 26 (June). Annual Reviews 4139 El Camino Way, PO Box 10139, Palo Alto, California 94303-0139, USA: 581–603.
Bhat, Neha, Hye-Joo Kwon, and Bruce B Riley. 2013. “A Gene Network That Coordinates Preplacodal Competence and Neural Crest Specification in Zebrafish.” *Developmental Biology* 373 (1): 107–17.
Bixby, Suzanne, Genevieve M Kruger, Jack T Mosher, Nancy M Joseph, and Sean J Morrison. 2002. “Cell-Intrinsic Differences between Stem Cells from Different Regions of the Peripheral Nervous System Regulate the Generation of Neural Diversity.” *Neuron* 35 (4): 643–56.
Black, J. 1989. “Drugs from Emasculated Hormones: The Principle of Syntopic Antagonism.” *Science (New York, N.Y.)* 245 (4917): 486–93.
Bockman, D E, and M L Kirby. 1984. “Dependence of Thymus Development on Derivatives of the Neural Crest.” *Science (New York, N.Y.)* 223 (4635): 498–500.
Boiani, Michele, and Hans R Schöler. 2005. “Regulatory Networks in Embryo-Derived Pluripotent Stem Cells.” *Nature Reviews Molecular Cell Biology* 6 (11). Nature Publishing Group: 872–84.
Boissy, Raymond E. 2003. “Melanosome Transfer to and Translocation in the Keratinocyte.” *Experimental Dermatology* 12 Suppl 2: 5–12.
Bondurand, N, V Pingault, D E Goerich, N Lemort, E Sock, C Le Caignec, M Wegner, and M Goossens. 2000. “Interaction among SOX10, PAX3 and MITF, Three Genes Altered in Waardenburg Syndrome.” *Human Molecular Genetics* 9 (13): 1907–17.
Breacker, C., and C. A. Tilley. 2014. “Promotion of a Non-Invasive DNA Sampling Technique of Zebrafish (Danio Rerio) and Sticklebacks (Gasterosteus Aculeatus).”
Britsch, S, D E Goerich, D Riethmacher, R I Peirano, M Rossner, K A Nave, C

Birchmeier, and M Wegner. 2001. “The Transcription Factor Sox10 Is a Key Regulator of Peripheral Glial Development.” *Genes & Development* 15 (1). Cold Spring Harbor Laboratory Press: 66–78.
Bronner-Fraser, M, and S E Fraser. 1988. “Cell Lineage Analysis Reveals Multipotency of Some Avian Neural Crest Cells.” *Nature* 335 (6186): 161–64.
Bronner-Fraser, Marianne, and Scott Fraser. 1989. “Developmental Potential of Avian Trunk Neural Crest Cells in Situ.” *Neuron* 3 (6): 755–66.
Bronner-Fraser, Marianne, and Scott E Fraser. 1991. “Cell Lineage Analysis of the Avian Neural Crest.” *Development Supplement* 2: 17–22.
Bruno, JoAnne. 1997. “Ligand Fishing with Biacore.” New Jersey.
Burda, Z, A Krzywicki, O C Martin, and M Zagorski. 2011. “Motifs Emerge from Function in Model Gene Regulatory Networks.” *Proceedings of the National Academy of Sciences of the United States of America* 108 (42): 17263–68. http://www.pnas.org/content/108/42/17263.abstract.
Burns, Alan J, Jean-Marie M Delalande, and Nicole M Le Douarin. 2002. “In Ovo Transplantation of Enteric Nervous System Precursors from Vagal to Sacral Neural Crest Results in Extensive Hindgut Colonisation.” *Development (Cambridge, England)* 129 (12). The Company of Biologists Ltd: 2785–96.
Carney, Thomas J, Kirsten A Dutton, Emma Greenhill, Mariana Delfino-Machín, Pascale Dufourcq, Patrick Blader, and Robert N Kelsh. 2006. “A Direct Role for Sox10 in Specification of Neural Crest-Derived Sensory Neurons.” *Development (Cambridge, England)* 133 (23): 4619–30.
Chai, Y, X Jiang, Y Ito, P Bringas, J Han, D H Rowitch, P Soriano, A P McMahon, and H M Sucov. 2000. “Fate of the Mammalian Cranial Neural Crest during Tooth and Mandibular Morphogenesis.” *Development (Cambridge, England)* 127 (8): 1671–79.
Chan, Yun Shen, Lin Yang, and Huck-Hui Ng. 2011. “Transcriptional Regulatory Networks in Embryonic Stem Cells.” *Progress in Drug Research.* 67 (January): 239–52.
Cheng, Y, M Cheung, M M Abu-Elmagd, A Orme, and P J Scotting. 2000. “Chick sox10, a Transcription Factor Expressed in Both Early Neural Crest Cells and Central Nervous System.” *Brain Research. Developmental Brain Research* 121 (2): 233–41.
Cheung, Martin, and James Briscoe. 2003. “Neural Crest Development Is Regulated by the Transcription Factor Sox9.” *Development (Cambridge, England)* 130 (23). The Company of Biologists Ltd: 5681–93.
Cheung, Martin, Marie-Christine Chaboissier, Anita Mynett, Elizabeth Hirst, Andreas Schedl, and James Briscoe. 2005. “The Transcriptional Control of Trunk Neural Crest Induction, Survival, and Delamination.” *Developmental Cell* 8 (2). Elsevier: 179–92.
Chomczynski, Piotr, and Nicoletta Sacchi. 1987. “Single-Step Method of RNA Isolation by Acid Guanidinium Thiocyanate-Phenol-Chloroform Extraction.” *Analytical Biochemistry* 162 (1). Academic Press: 156–59. http://linkinghub.elsevier.com/retrieve/pii/0003269787900212.
Chong, Shang-Wei, Thi-Thu-Hang Nguyen, Lee-Thean Chu, Yun-Jin Jiang, and Vladimir Korzh. 2005. “Zebrafish id2 Developmental Expression Pattern Contains Evolutionary Conserved and Species-Specific Characteristics.” *Developmental Dynamics* 234 (4). Wiley-Liss, Inc.: 1055–63.
Chuang, Han-Ni, Hsiao-Yun Cheng, Kuang-Ming Hsiao, Chieh-Wen Lin, Mei-Ling Lin, and Huichin Pan. 2010. “The Zebrafish Homeobox Gene irxl1 Is Required for Brain and Pharyngeal Arch Morphogenesis.” *Developmental Dynamics* 239 (2): 639–50.
Cirillo, Lisa Ann, Frank Robert Lin, Isabel Cuesta, Dara Friedman, Michal Jarnik, and Kenneth S Zaret. 2002. “Opening of Compacted Chromatin by Early Developmental Transcription Factors HNF3 (FoxA) and GATA-4.” *Molecular Cell* 9 (2): 279–89.
Clark, Kirk L., Elaine D Halay, Eseng Lai, and Stephen K Burley. 1993. “Co-Crystal

Structure of the HNF-3/fork Head DNA-Recognition Motif Resembles Histone H5." *Nature* 364 (6436). Nature Publishing Group: 412–20.
Cloney, Richard A, and Steven L Brocco. 1983. "Chromatophore Organs, Reflector Cells, Iridocytes and Leucophores in Cephalopods." *American Zoologist* 23 (3). Soc Integ Comp Biol: 581–92.
Cohen, Michael, Anna Kicheva, Ana Ribeiro, Robert Blassberg, Karen M Page, Chris P Barnes, and James Briscoe. 2015. "Ptch1 and Gli Regulate Shh Signalling Dynamics via Multiple Mechanisms." *Nature Communications* 6: 6709.
Cohen, Michael, Karen M Page, Ruben Perez-Carrasco, Chris P Barnes, and James Briscoe. 2014. "A Theoretical Framework for the Regulation of Shh Morphogen-Controlled Gene Expression." *Development (Cambridge, England)* 141 (20): 3868–78.
Collazo, A, M Bronner-Fraser, and S E Fraser. 1993. "Vital Dye Labelling of Xenopus Laevis Trunk Neural Crest Reveals Multipotency and Novel Pathways of Migration." *Development* 118 (2): 363–76.
Connor, F, P D Cary, C M Read, N S Preston, P C Driscoll, P Denny, C Crane-Robinson, and A Ashworth. 1994. "DNA Binding and Bending Properties of the Post-Meiotically Expressed Sry-Related Protein Sox-5." *Nucleic Acids Research* 22 (16): 3339–46.
Cooper, Cynthia D, Tor H Linbo, and David W Raible. 2009. "Kit and foxd3 Genetically Interact to Regulate Melanophore Survival in Zebrafish." *Developmental Dynamics* 238 (4): 875–86.
Crespo, Isaac, and Antonio del Sol. 2013. "A General Strategy for Cellular Reprogramming: The Importance of Transcription Factor Cross-Repression." *Stem Cells* 31 (10): 2127–35.
Creuzet, Sophie E, Salvador Martinez, and Nicole M Le Douarin. 2006. "The Cephalic Neural Crest Exerts a Critical Effect on Forebrain and Midbrain Development." *Proceedings of the National Academy of Sciences of the United States of America* 103 (38). National Academy of Sciences: 14033–38.
Crombach, Anton, Karl R. Wotton, Damjan Cicin-Sain, Maksat Ashyraliyev, and Johannes Jaeger. 2012. "Efficient Reverse-Engineering of a Developmental Gene Regulatory Network." *PLoS Computational Biology*.
Crozé, Noémie de, Frédérique Maczkowiak, and Anne H Monsoro-Burq. 2011. "Reiterative AP2a Activity Controls Sequential Steps in the Neural Crest Gene Regulatory Network." *Proceedings of the National Academy of Sciences of the United States of America* 108 (1): 155–60.
Cuesta, Isabel, Kenneth S Zaret, and Pilar Santisteban. 2007. "The Forkhead Factor FoxE1 Binds to the Thyroperoxidase Promoter during Thyroid Cell Differentiation and Modifies Compacted Chromatin Structure." *Molecular and Cellular Biology* 27 (20). American Society for Microbiology (ASM): 7302–14.
Curran, Kevin, James A Lister, Gary R Kunkel, Andrew Prendergast, David M Parichy, and David W Raible. 2010. "Interplay between Foxd3 and Mitf Regulates Cell Fate Plasticity in the Zebrafish Neural Crest." *Developmental Biology* 344 (1). Elsevier Inc.: 107–18.
Curran, Kevin, David W Raible, and James A Lister. 2009. "Foxd3 Controls Melanophore Specification in the Zebrafish Neural Crest by Regulation of Mitf." *Developmental Biology* 332 (2). Elsevier Inc.: 408–17.
D'Mello, Stacey A N, Graeme J Finlay, Bruce C Baguley, and Marjan E Askarian-Amiri. 2016. "Signaling Pathways in Melanogenesis." *International Journal of Molecular Sciences* 17 (7). Multidisciplinary Digital Publishing Institute (MDPI).
Delfino-Machin, M, T R Chipperfield, Fslm Rodrigues, and R N Kelsh. 2007. "The Proliferating Field of Neural Crest Stem Cells." *Developmental Dynamics* 236 (12): 3242–54.
Dent, J A, A G Polson, and M W Klymkowsky. 1989. "A Whole-Mount Immunocytochemical Analysis of the Expression of the Intermediate Filament Protein Vimentin in Xenopus." *Development (Cambridge, England)* 105 (1): 61–74.

Dickmeis, Thomas, Sepand Rastegar, Chen Sok Lam, Pia Aanstad, Matthew Clark, Nadine Fischer, Frédéric Rosa, Vladimir Korzh, and Uwe Strähle. 2002. "Expression of the Helix-Loop-Helix Gene id3 in the Zebrafish Embryo." *Mechanisms of Development* 113 (1): 99–102.
Dietz, Marina S, Daniel Haße, Davide M Ferraris, Antonia Göhler, Hartmut H Niemann, and Mike Heilemann. 2013. "Single-Molecule Photobleaching Reveals Increased MET Receptor Dimerization upon Ligand Binding in Intact Cells." *BMC Biophysics* 6 (1): 6.
Diotel, Nicolas, Tanja Beil, Uwe Strähle, and Sepand Rastegar. 2015. "Differential Expression of Id Genes and Their Potential Regulator znf238 in Zebrafish Adult Neural Progenitor Cells and Neurons Suggests Distinct Functions in Adult Neurogenesis." *Gene Expression Patterns* 19 (1): 1–13.
Ditlevsen, Susanne, and Adeline Samson. 2013. "Introduction to Stochastic Models in Biology." In *Springer-Verlag Berlin Heidelberg*.
Dorsky, R I, R T Moon, and D W Raible. 1998. "Control of Neural Crest Cell Fate by the Wnt Signalling Pathway." *Nature* 396 (6709): 370–73.
Dorsky, Richard I, David W Raible, and Randall T Moon. 2000. "Direct Regulation of Nacre, a Zebrafish MITF Homolog Required for Pigment Cell Formation, by the Wnt Pathway." *Genes & Development* 14 (2). Cold Spring Harbor Laboratory Press: 158–62.
Dottori, M, M K Gross, P Labosky, and M Goulding. 2001. "The Winged-Helix Transcription Factor Foxd3 Suppresses Interneuron Differentiation and Promotes Neural Crest Cell Fate." *Development (Cambridge, England)* 128 (21): 4127–38.
Douarin, N M Le, and M A Teillet. 1973. "The Migration of Neural Crest Cells to the Wall of the Digestive Tract in Avian Embryo." *Journal of Embryology and Experimental Morphology* 30 (1). The Company of Biologists Ltd: 31–48.
Douarin, Nicole Le, and Chaya Kalcheim. 1999. *The Neural Crest*. Cambridge University Press.
Douarin, Nicole M Le, Giordano W Calloni, and Elisabeth Dupin. 2008. "The Stem Cells of the Neural Crest." *Cell Cycle* 7 (8). Landes Bioscience: 1013–19.
Douarin, Nicole M Le, and Elisabeth Dupin. 2003. "Multipotentiality of the Neural Crest." *Current Opinion in Genetics & Development* 13 (5): 529–36.
Dupin, Elisabeth, and Juliana M Coelho-Aguiar. 2013. "Isolation and Differentiation Properties of Neural Crest Stem Cells." *Cytometry. Part A : The Journal of the International Society for Analytical Cytology* 83 (1): 38–47.
Dupin, Elisabeth, Sophie Creuzet, and Nicole M. Le Douarin. 2006. "The Contribution of the Neural Crest to the Vertebrate Body." In , 96–119. Springer US.
Dutton, K A, A Pauliny, S S Lopes, S Elworthy, T J Carney, J Rauch, R Geisler, P Haffter, and R N Kelsh. 2001. "Zebrafish Colourless Encodes sox10 and Specifies Non-Ectomesenchymal Neural Crest Fates." *Development* 128 (21): 4113–25.
Eden, Eran, Naama Geva-Zatorsky, Irina Issaeva, Ariel Cohen, Erez Dekel, Tamar Danon, Lydia Cohen, Avi Mayo, and Uri Alon. 2011. "Proteome Half-Life Dynamics in Living Human Cells." *Science (New York, N.Y.)* 331 (6018): 764–68.
Eisen, Judith S, and James C Smith. 2008. "Controlling Morpholino Experiments: Don't Stop Making Antisense." *Development* 135 (10): 1735–43.
Elworthy, Stone, James A Lister, Tom J Carney, David W Raible, and Robert N Kelsh. 2003. "Transcriptional Regulation of Mitfa Accounts for the sox10 Requirement in Zebrafish Melanophore Development." *Development* 130 (12): 2809–18.
Epstein, M. L., T. Mikawa, A. M. C. Brown, and D. R. McFarlin. 1994. "Mapping the Origin of the Avian Enteric Nervous System with a Retroviral Marker." *Developmental Dynamics* 201 (3). Wiley Subscription Services, Inc., A Wiley Company: 236–44.
Erickson, C A, and M V Reedy. 1998. "Neural Crest Development: The Interplay between Morphogenesis and Cell Differentiation." *Current Topics in Developmental Biology* 40: 177–209.
Fadeev, Andrey, Jana Krauss, Ajeet Pratap Singh, and Christiane Nüsslein-Volhard. 2016. "Zebrafish Leucocyte Tyrosine Kinase Controls Iridophore Establishment,

Proliferation and Survival." *Pigment Cell & Melanoma Research* 29 (3): 284–96.
Farell, Eric M, and Gladys Alexandre. 2012. "Bovine Serum Albumin Further Enhances the Effects of Organic Solvents on Increased Yield of Polymerase Chain Reaction of GC-Rich Templates." *BMC Research Notes* 5: 257.
Fomekong-Nanfack, Yves, Marten Postma, Jaap A Kaandorp, E Davidson, H Kitano, CC Guet, MB Elowitz, et al. 2009. "Inferring Drosophila Gap Gene Regulatory Network: A Parameter Sensitivity and Perturbation Analysis." *BMC Systems Biology* 3 (1). BioMed Central: 94.
Fox, Denis L. 1976. *Animal Biochromes and Structural Colours : Physical, Chemical, Distributional and Physiological Features of Coloured Bodies in the Animal World*. University of California Press.
Frank, E, and J R Sanes. 1991. "Lineage of Neurons and Glia in Chick Dorsal Root Ganglia: Analysis in Vivo with a Recombinant Retrovirus." *Development (Cambridge, England)* 111 (4): 895–908.
Furey, Terrence S. 2012. "ChIP-Seq and beyond: New and Improved Methodologies to Detect and Characterize Protein-DNA Interactions." *Nature Reviews Genetics* 13 (12). Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.: 840–52.
Gaj, Thomas, Charles A Gersbach, Carlos F Barbas, and III. 2013. "ZFN, TALEN, and CRISPR/Cas-Based Methods for Genome Engineering." *Trends in Biotechnology* 31 (7). NIH Public Access: 397–405.
Gammill, Laura S., and Marianne Bronner-Fraser. 2003. "Neural Crest Specification: Migrating into Genomics." *Nature Reviews Neuroscience* 4 (10): 795–805.
Gans, Carl, and R. Glenn Northcutt. 1983. "Neural Crest and the Origin of Vertebrates: A New Head." *Science* 220 (4594). American Association for the Advancement of Science: 221–26.
Garnett, Aaron T, Tyler A Square, Daniel M Medeiros, A. G. Bang, N. Papalopulu, M. D. Goulding, C. Kintner, et al. 2012. "BMP, Wnt and FGF Signals Are Integrated through Evolutionarily Conserved Enhancers to Achieve Robust Expression of Pax3 and Zic Genes at the Zebrafish Neural Plate Border." *Development (Cambridge, England)* 139 (22). Oxford University Press for The Company of Biologists Limited: 4220–31.
Gilbert, S. F. 2000. "The Neural Crest." In *Developmental Biology*, 6th ed. Sunderland (MA): Sinauer Associates.
Gilbert, Scott. F. 2013. *Developmental Biology*. 10th ed. Sinauer Associates.
Gillespie, Daniel T. 1977. "Exact Stochastic Simulation of Coupled Chemical Reactions." *Journal of Physical Chemistry*.
Glavic, Alvaro, Francisca Silva, Manuel J Aybar, Francisco Bastidas, and Roberto Mayor. 2004. "Interplay between Notch Signaling and the Homeoprotein Xiro1 Is Required for Neural Crest Induction in Xenopus Embryos." *Development (Cambridge, England)* 131 (2): 347–59.
Goda, Makoto, and Ryozo Fujii. 1995. "Blue Chromatophores in Two Species of Callionymid Fish." *Zoological Science* 12 (6): 811–13.
Goda, Makoto, Mihoko Ohata, Hayato Ikoma, Yoshinori Fujiyoshi, Masazumi Sugimoto, and Ryozo Fujii. 2011. "Integumental Reddish-Violet Coloration Owing to Novel Dichromatic Chromatophores in the Teleost Fish, Pseudochromis Diadema." *Pigment Cell & Melanoma Research* 24 (4): 614–17.
Gorkin, David U, Dongwon Lee, Xylena Reed, Christopher Fletez-Brant, Seneca L Bessling, Stacie K Loftus, Michael A Beer, William J Pavan, and Andrew S McCallion. 2012. "Integration of ChIP-Seq and Machine Learning Reveals Enhancers and a Predictive Regulatory Sequence Vocabulary in Melanocytes." *Genome Research* 22 (11). Cold Spring Harbor Laboratory Press: 2290–2301.
Graham, Anthony, Jo Begbie, and Imelda McGonnell. 2004. "Significance of the Cranial Neural Crest." *Developmental Dynamics : An Official Publication of the American Association of Anatomists* 229 (1): 5–13.
Greenhill, Emma R, Andrea Rocco, Laura Vibert, Masataka Nikaido, and Robert N Kelsh. 2011. "An Iterative Genetic and Dynamical Modelling Approach Identifies

Novel Features of the Gene Regulatory Network Underlying Melanocyte Development." Edited by Mary C Mullins. *PLoS Genetics* 7 (9). Public Library of Science: 18.
Griffiths, Anthony JF, Jeffrey H Miller, David T Suzuki, Richard C Lewontin, and William M Gelbart. 2000. "Chi-Square Test." W. H. Freeman.
Gross-Thebing, Theresa, Azadeh Paksa, Erez Raz, G Hauptmann, T Gerster, C Thisse, B Thisse, et al. 2014. "Simultaneous High-Resolution Detection of Multiple Transcripts Combined with Localization of Proteins in Whole-Mount Embryos." *BMC Biology* 12 (1). BioMed Central: 55.
Groves, Andrew K., and Carole LaBonne. 2014. "Setting Appropriate Boundaries: Fate, Patterning and Competence at the Neural Plate Border." *Developmental Biology* 389 (1): 2–12.
Guan, Jikui, Ganesh Umapathy, Yasuo Yamazaki, Georg Wolfstetter, Patricia Mendoza, Kathrin Pfeifer, Ateequrrahman Mohammed, et al. 2015. "FAM150A and FAM150B Are Activating Ligands for Anaplastic Lymphoma Kinase." *eLife* 4. eLife Sciences Publications, Ltd: e09811.
Gunawardena, J. 2014. "Models in Biology: 'accurate Descriptions of Our Pathetic Thinking.'" *BMC Biology* 12 (1). BioMed Central: 29.
Hall, B. K. 2008. "The Neural Crest and Neural Crest Cells: Discovery and Significance for Theories of Embryonic Organization." *Journal of Biosciences* 33 (5). Springer: 781–93.
Hall, B K. 2000. "The Neural Crest as a Fourth Germ Layer and Vertebrates as Quadroblastic Not Triploblastic." *Evolution & Development* 2 (1): 3–5.
Hall, Brian K. 2008. "The Neural Crest and Neural Crest Cells: Discovery and Significance for Theories of Embryonic Organization." *Journal of Biosciences* 33 (5): 781–93.
Hallsson, Jón H, Benedikta S Haflidadóttir, Chad Stivers, Ward Odenwald, Heinz Arnheiter, Francesca Pignoni, Eiríkur Steingrímsson, et al. 2004. "The Basic Helix-Loop-Helix Leucine Zipper Transcription Factor Mitf Is Conserved in Drosophila and Functions in Eye Development." *Genetics* 167 (1). Genetics: 233–41.
Hanna, Lynn A, Ruth K Foreman, Illya A Tarasenko, Daniel S Kessler, and Patricia A Labosky. 2002. "Requirement for Foxd3 in Maintaining Pluripotent Cells of the Early Mouse Embryo." *Genes & Development* 16 (20): 2650–61.
Hans, Stefan, Joe Christison, Dong Liu, Monte Westerfield, DM Noden, TR van de Water, GF Couly, et al. 2007. "Fgf-Dependent Otic Induction Requires Competence Provided by Foxi1 and Dlx3b." *BMC Developmental Biology* 7 (1). BioMed Central: 5.
Harris, Arthur J. 1912. "A Simple Test of the Goodness of Fit of Mendelian Ratios." *The American Naturalist* 46 (552): 741–45. http://www.jstor.org/stable/2456130.
Hatta, Mitsutoki, and Lisa Ann Cirillo. 2007. "Chromatin Opening and Stable Perturbation of Core histone:DNA Contacts by FoxO1." *The Journal of Biological Chemistry* 282 (49): 35583–93.
Hauser, Stefan, Darius Widera, Firas Qunneis, Janine Müller, Christin Zander, Johannes Greiner, Christina Strauss, et al. 2012. "Isolation of Novel Multipotent Neural Crest-Derived Stem Cells from Adult Human Inferior Turbinate." *Stem Cells and Development* 21 (5): 742–56.
He, T C, A B Sparks, C Rago, H Hermeking, L Zawel, L T da Costa, P J Morin, et al. 1998. "Identification of c-MYC as a Target of the APC Pathway." *Science (New York, N.Y.)* 281 (5382). American Association for the Advancement of Science: 1509–12.
Heanue, Tiffany A., Iain T. Shepherd, and Alan J. Burns. 2016. "Enteric Nervous System Development in Avian and Zebrafish Models." *Developmental Biology*.
Hearing, Vincent J. 2011. "Determination of Melanin Synthetic Pathways." *J Invest Dermatol* 17 (131).
Hegna, Robert H, Ossi Nokelainen, Jonathan R Hegna, Johanna Mappes, SM. Gray, JS. McKinnon, JR. True, et al. 2013. "To Quiver or to Shiver: Increased Melanization Benefits Thermoregulation, but Reduces Warning Signal Efficacy in

the Wood Tiger Moth." *Proceedings. Biological Sciences / The Royal Society* 280 (1755). The Royal Society: 20122812.
Hemesath, T J, E Steingrímsson, G McGill, M J Hansen, J Vaught, C A Hodgkinson, H Arnheiter, N G Copeland, N A Jenkins, and D E Fisher. 1994. "Microphthalmia, a Critical Factor in Melanocyte Development, Defines a Discrete Transcription Factor Family." *Genes & Development* 8 (22): 2770–80.
Henion, P D, and J A Weston. 1997. "Timing and Pattern of Cell Fate Restrictions in the Neural Crest Lineage." *Development* 124 (21): 4351–59.
Hernandez-Lagunas, Laura, Irene F Choi, Takao Kaji, Peter Simpson, Candice Hershey, Yi Zhou, Len Zon, Mark Mercola, and Kristin Bruk Artinger. 2005. "Zebrafish Narrowminded Disrupts the Transcription Factor prdm1 and Is Required for Neural Crest and Sensory Neuron Specification." *Developmental Biology* 278 (2). NIH Public Access: 347–57.
Higdon, Charles W, Robi D Mitra, and Stephen L Johnson. 2013. "Gene Expression Analysis of Zebrafish Melanocytes, Iridophores, and Retinal Pigmented Epithelium Reveals Indicators of Biological Function and Developmental Origin." *PLoS One* 8 (7): e67801.
Hirschi, Karen K, and Mark W Majesky. 2004. "Smooth Muscle Stem Cells." *The Anatomical Record. Part A, Discoveries in Molecular, Cellular, and Evolutionary Biology* 276 (1): 22–33.
Hochgreb-Hägele, Tatiana, and Marianne E. Bronner. 2013. "A Novel FoxD3 Gene Trap Line Reveals Neural Crest Precursor Movement and a Role for FoxD3 in Their Specification." *Developmental Biology* 374 (1): 1–11.
Hoffman, Trevor L., Anna L. Javier, Shelley A. Campeau, Robert D. Knight, and Thomas F. Schilling. 2007. "Tfap2 Transcription Factors in Zebrafish Neural Crest Development and Ectodermal Evolution." *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution* 308B (5). Wiley Subscription Services, Inc., A Wiley Company: 679–91.
Holmberg, Johan, and Thomas Perlmann. 2012. "Maintaining Differentiated Cellular Identity." *Nature Reviews Genetics* 13 (6). Nature Publishing Group: 429–39.
Hong, C.-S., and J.-P. Saint-Jeannet. 2007. "The Activity of Pax3 and Zic1 Regulates Three Distinct Cell Fates at the Neural Plate Border." *Molecular Biology of the Cell* 18 (6). American Society for Cell Biology: 2192–2202.
Hong, Chang-soo. 2005. "Sox Proteins and Neural Crest Development." *Seminars in Cell Developmental Biology* 16 (6): 694–703.
Honoré, Stella M, Manuel J Aybar, and Roberto Mayor. 2003. "Sox10 Is Required for the Early Development of the Prospective Neural Crest in Xenopus Embryos." *Developmental Biology* 260 (1): 79–96.
Hoshijima, Kazuyuki, Michael J. Jurynec, and David Jonah Grunwald. 2016. "Precise Editing of the Zebrafish Genome Made Simple and Efficient." *Developmental Cell* 36 (6): 654–67.
Hosoda, K, R E Hammer, J A Richardson, A G Baynash, J C Cheung, A Giaid, and M Yanagisawa. 1994. "Targeted and Natural (Piebald-Lethal) Mutations of Endothelin-B Receptor Gene Produce Megacolon Associated with Spotted Coat Color in Mice." *Cell* 79 (7): 1267–76.
Huang, E J, K Zang, A Schmidt, A Saulys, M Xiang, and L F Reichardt. 1999. "POU Domain Factor Brn-3a Controls the Differentiation and Survival of Trigeminal Neurons by Regulating Trk Receptor Expression." *Development (Cambridge, England)* 126 (13): 2869–82.
Huber, Katrin. 2006. "The Sympathoadrenal Cell Lineage: Specification, Diversification, and New Perspectives." *Developmental Biology* 298 (2): 335–43.
Hwang, Woong Y, Yanfang Fu, Deepak Reyon, Morgan L Maeder, Shengdar Q Tsai, Jeffry D Sander, Randall T Peterson, J-R Joanna Yeh, and J Keith Joung. 2013. "Efficient Genome Editing in Zebrafish Using a CRISPR-Cas System." *Nature Biotechnology* 31 (3). Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.: 227–29.
Ignatius, Myron S, Holly E Moose, Heithem M El-Hodiri, and Paul D Henion. 2008.

"colgate/hdac1 Repression of foxd3 Expression Is Required to Permit Mitfa-Dependent Melanogenesis." *Developmental Biology* 313 (2): 568–83.
Izumi, K, T Kohta, Y Kimura, S Ishida, T Takahashi, A Ishiko, and K Kosaki. 2008. "Tietz Syndrome: Unique Phenotype Specific to Mutations of MITF Nuclear Localization Signal." *Clinical Genetics* 74 (1): 93–95.
Janebodin, Kajohnkiart, Orapin V Horst, Nicholas Ieronimakis, Gayathri Balasundaram, Kanit Reesukumal, Busadee Pratumvinit, and Morayma Reyes. 2011. "Isolation and Characterization of Neural Crest-Derived Stem Cells from Dental Pulp of Neonatal Mice." *PloS One* 6 (11): e27526.
Jones, J. 2008. "Table: Chi-Square Probabilities." http://people.richland.edu/james/lecture/m170/tbl-chi.html.
Jung, Sin-Ho, and Insuk Sohn. 2014. "Statistical Issues in the Design and Analysis of nCounter Projects." *Cancer Informatics* 13 (Suppl 7). Libertas Academica: 35–43.
Kadono, S, I Manaka, M Kawashima, T Kobayashi, and G Imokawa. 2001. "The Role of the Epidermal Endothelin Cascade in the Hyperpigmentation Mechanism of Lentigo Senilis." *The Journal of Investigative Dermatology* 116 (4): 571–77.
Kague, Erika, Michael Gallagher, Sally Burke, Michael Parsons, Tamara Franz-Odendaal, and Shannon Fisher. 2012. "Skeletogenic Fate of Zebrafish Cranial and Trunk Neural Crest." *PLoS One* 7 (11): e47394.
Kaufman, Charles K., Christian Mosimann, Zi Peng Fan, Song Yang, Andrew J. Thomas, Julien Ablain, Justin L. Tan, et al. 2016. "A Zebrafish Melanoma Model Reveals Emergence of Neural Crest Identity during Melanoma Initiation." *Science* 351 (6272).
Kelsh, R.N. 2006. "Sorting out Sox10 Functions in Neural Crest Development." *BioEssays News and Reviews in Molecular Cellular and Developmental Biology* 28 (8). Wiley Subscription Services, Inc., A Wiley Company: 788–98.
Kelsh, R N, M Brand, Y J Jiang, C P Heisenberg, S Lin, P Haffter, J Odenthal, et al. 1996. "Zebrafish Pigmentation Mutations and the Processes of Neural Crest Development." *Development* 123 (1): 369–89.
Kelsh, R N, and J S Eisen. 2000. "The Zebrafish Colourless Gene Regulates Development of Non-Ectomesenchymal Neural Crest Derivatives." *Development* 127 (3): 2581–90.
Kelsh, R N, B Schmid, and J S Eisen. 2000. "Genetic Analysis of Melanophore Development in Zebrafish Embryos." *Developmental Biology* 225 (2): 277–93.
Kelsh, RN, and CA Erickson. 2013. *Neural Crest: Origin, Migration and Differentiation*. John Wiley & Sons, Ltd.
Kelsh, Robert N. 2004. "Genetics and Evolution of Pigment Patterns in Fish." *Pigment Cell & Melanoma Research* 17 (4): 326–36.
Kelsh, Robert N, Melissa L Harris, Sarah Colanesi, and Carol A Erickson. 2009. "Stripes and Belly-Spots -- a Review of Pigment Cell Morphogenesis in Vertebrates." *Seminars in Cell Developmental Biology* 20 (1). Elsevier Ltd: 90–104.
Khaliq, Mehwish, Tae-Young Choi, Juhoon So, and Donghun Shin. 2015. "Id2a Is Required for Hepatic Outgrowth during Liver Development in Zebrafish." *Mechanisms of Development* 138 Pt 3 (November): 399–414.
Khudyakov, Jane, and Marianne Bronner-Fraser. 2009. "Comprehensive Spatiotemporal Analysis of Early Chick Neural Crest Network Genes." *Developmental Dynamics : An Official Publication of the American Association of Anatomists* 238 (3): 716–23.
Kidder, Benjamin L, Gangqing Hu, and Keji Zhao. 2011. "ChIP-Seq: Technical Considerations for Obtaining High-Quality Data." *Nature Immunology* 12 (10). NIH Public Access: 918–22.
Kimelman, David. 2006. "Mesoderm Induction: From Caps to Chips." *Nature Reviews Genetics* 7 (5). Nature Publishing Group: 360–72.
Kimmel, C B, W W Ballard, S R Kimmel, B Ullmann, and T F Schilling. 1995. "Stages of Embryonic Development of the Zebrafish." *Developmental Dynamics an Official Publication of the American Association of Anatomists* 203 (3). Wiley Online

Library: 253–310.
Kimura, Yasuyoshi, Masaaki Oda, Tsunetoshi Nakatani, Yoichi Sekita, Asun Monfort, Anton Wutz, Hideki Mochizuki, and Toru Nakano. 2015. “CRISPR/Cas9-Mediated Reporter Knock-in in Mouse Haploid Embryonic Stem Cells.” *Scientific Reports* 5: 10710.
Kirby, M L, T F Gale, and D E Stewart. 1983. “Neural Crest Cells Contribute to Normal Aorticopulmonary Septation.” *Science (New York, N.Y.)* 220 (4601): 1059–61.
Kirk, P. D. W., A. C. Babtie, and M. P. H. Stumpf. 2015. “Systems Biology (Un)certainties.” *Science* 350 (6259).
Knight, Robert D., and Thomas F Schilling. 2013. “Cranial Neural Crest and Development of the Head Skeleton.” Landes Bioscience.
Knight, Robert D, Yashar Javidan, Sarah Nelson, Tailin Zhang, and Thomas Schilling. 2004. “Skeletal and Pigment Cell Defects in the Lockjaw Mutant Reveal Multiple Roles for Zebrafish tfap2a in Neural Crest Development.” *Developmental Dynamics* 229 (1): 87–98.
Kolodziej, Katarzyna E, Farzin Pourfarzad, Ernie de Boer, Sanja Krpic, Frank Grosveld, and John Strouboulis. 2009. “Optimal Use of Tandem Biotin and V5 Tags in ChIP Assays.” *BMC Molecular Biology* 10: 6.
Kopfova, J. 2006. “Hysteresis in Biological Models.” *Journal of Physics: Conference Series* 55: 130–34.
Kormish, Jay D, Débora Sinner, and Aaron M Zorn. 2010. “Interactions between SOX Factors and Wnt/beta-Catenin Signaling in Development and Disease.” *Developmental Dynamics : An Official Publication of the American Association of Anatomists* 239 (1): 56–68.
Kos, R, M V Reedy, R L Johnson, and C A Erickson. 2001. “The Winged-Helix Transcription Factor FoxD3 Is Important for Establishing the Neural Crest Lineage and Repressing Melanogenesis in Avian Embryos.” *Development* 128 (8): 1467–79.
Kozlosky, C J, E Maraskovsky, J T McGrew, T VandenBos, M Teepe, S D Lyman, S Srinivasan, F A Fletcher, R B Gayle, and D P Cerretti. 1995. “Ligands for the Receptor Tyrosine Kinases Hek and Elk: Isolation of cDNAs Encoding a Family of Proteins.” *Oncogene* 10 (2): 299–306.
Krauss, Jana, Pantelis Astrinidis, Pantilis Astrinides, Hans Georg Frohnhöfer, Brigitte Walderich, and Christiane Nüsslein-Volhard. 2013. “Transparent, a Gene Affecting Stripe Formation in Zebrafish, Encodes the Mitochondrial Protein Mpv17 That Is Required for Iridophore Survival.” *Biology Open* 2 (7). Company of Biologists: 703–10.
Krauss, Jana, Hans Georg Frohnhöfer, Brigitte Walderich, Hans-Martin Maischein, Christian Weiler, Uwe Irion, and Christiane Nüsslein-Volhard. 2014. “Endothelin Signalling in Iridophore Development and Stripe Pattern Formation of Zebrafish.” *Biology Open* 3 (6): 503–9.
Krishnakumar, Raga, Amy F. Chen, Marisol G. Pantovich, Muhammad Danial, Ronald J. Parchem, Patricia A. Labosky, and Robert Blelloch. 2016. “FOXD3 Regulates Pluripotent Stem Cell Potential by Simultaneously Initiating and Repressing Enhancer Activity.” *Cell Stem Cell* 18 (1): 104–17.
Krispin, Shlomo, Erez Nitzan, Yachia Kassem, and Chaya Kalcheim. 2010. “Evidence for a Dynamic Spatiotemporal Fate Map and Early Fate Restrictions of Premigratory Avian Neural Crest.” *Development* 137 (4): 585–95.
Kroll, K L, A N Salic, L M Evans, and M W Kirschner. 1998. “Geminin, a Neuralizing Molecule That Demarcates the Future Neural Plate at the Onset of Gastrulation.” *Development (Cambridge, England)* 125 (16): 3247–58.
Kruger, Genevieve M., Jack T. Mosher, Suzanne Bixby, Nancy Joseph, Toshihide Iwashita, Sean J. Morrison, J. Altman, et al. 2002. “Neural Crest Stem Cells Persist in the Adult Gut but Undergo Changes in Self-Renewal, Neuronal Subtype Potential, and Factor Responsiveness.” *Neuron* 35 (4). Elsevier: 657–69.
Krumsiek, Jan, Carsten Marr, Timm Schroeder, and Fabian J Theis. 2011. “Hierarchical Differentiation of Myeloid Progenitors Is Encoded in the Transcription

Factor Network." Edited by Maurizio Pesce. *PLoS One* 6 (8). Public Library of Science: e22649.
Kudoh, Tetsuhiro, Stephen W Wilson, and Igor B Dawid. 2002. "Distinct Roles for Fgf, Wnt and Retinoic Acid in Posteriorizing the Neural Ectoderm." *Development (Cambridge, England)* 129 (18): 4335–46.
Kueh, Hao Yuan, and Ellen V Rothenberg. 2013. "Regulatory Gene Network Circuits Underlying T Cell Development from Multipotent Progenitors." *Wiley Interdisciplinary Reviews. Systems Biology and Medicine* 4 (1): 79–102.
Kuhar, Michael J. 2010. "Measuring Levels of Proteins by Various Technologies: Can We Learn More by Measuring Turnover?" *Biochemical Pharmacology* 79 (5): 665–68.
Kuiper, Roland P, Marga Schepens, José Thijssen, Eric F P M Schoenmakers, and Ad Geurts van Kessel. 2004. "Regulation of the MiTF/TFE bHLH-LZ Transcription Factors through Restricted Spatial Expression and Alternative Splicing of Functional Domains." *Nucleic Acids Research* 32 (8): 2315–22.
Kulesa, Paul M, Jennifer C Kasemeier-Kulesa, Jessica M Teddy, Naira V Margaryan, Elisabeth A Seftor, Richard E B Seftor, and Mary J C Hendrix. 2006. "Reprogramming Metastatic Melanoma Cells to Assume a Neural Crest Cell-like Phenotype in an Embryonic Microenvironment." *Proceedings of the National Academy of Sciences of the United States of America* 103 (10). National Academy of Sciences: 3752–57.
Kwak, Jina, Ok Kyu Park, Yoo Jung Jung, Byung Joon Hwang, Seung-Hae Kwon, and Yun Kee. 2013. "Live Image Profiling of Neural Crest Lineages in Zebrafish Transgenic Lines." *Molecules and Cells* 35 (3): 255–60.
Kwon, Hye-Joo, Neha Bhat, Elly M Sweet, Robert A Cornell, and Bruce B Riley. 2010. "Identification of Early Requirements for Preplacodal Ectoderm and Sensory Organ Development." *PLoS Genetics* 6 (9): e1001133.
Lackmann, M, T Bucci, R J Mann, L A Kravets, E Viney, F Smith, R L Moritz, et al. 1996. "Purification of a Ligand for the EPH-like Receptor HEK Using a Biosensor-Based Affinity Detection Approach." *Proceedings of the National Academy of Sciences of the United States of America* 93 (6): 2523–27.
Land, M F. 1972. "The Physics and Biology of Animal Reflectors." *Progress in Biophysics and Molecular Biology* 24: 75–106.
Ledda, Fernanda, and Gustavo Paratcha. 2007. "Negative Regulation of Receptor Tyrosine Kinase (RTK) Signaling: A Developing Field." *Biomarker Insights* 2. Libertas Academica: 45–58.
Lee, Hung-Chieh, Hsing-Yen Huang, Cheng-Yung Lin, Yau-Hung Chen, and Huai-Jen Tsai. 2006. "Foxd3 Mediates Zebrafish myf5 Expression during Early Somitogenesis." *Developmental Biology* 290 (2): 359–72.
Lee, Hwa Jin, Brian Wall, and Suzie Chen. 2008. "G-Protein-Coupled Receptors and Melanoma." *Pigment Cell & Melanoma Research* 21 (4): 415–28.
Lei, Lei, Jing Zhou, Lu Lin, and Luis F. Parada. 2006. "Brn3a and Klf7 Cooperate to Control TrkA Expression in Sensory Neurons." *Developmental Biology* 300 (2): 758–69.
Lemke, Greg. 2015. "Adopting ALK and LTK." *Proceedings of the National Academy of Sciences of the United States of America* 112 (52). National Academy of Sciences: 15783–84.
Leung, Anskar Y H, Eric M Mendenhall, Tommy T F Kwan, Raymond Liang, Craig Eckfeldt, Eleanor Chen, Matthias Hammerschmidt, Suzanne Grindley, Stephen C Ekker, and Catherine M Verfaillie. 2005. "Characterization of Expanded Intermediate Cell Mass in Zebrafish Chordin Morphant Embryos." *Developmental Biology* 277 (1): 235–54.
Leung, Chuen Yan, and Magdalena Zernicka-Goetz. 2015. "Mapping the Journey from Totipotency to Lineage Specification in the Mouse Embryo." *Current Opinion in Genetics & Development* 34: 71–76.
Levine, Michael, and Eric H Davidson. 2005. "Gene Regulatory Networks for Development." *Proceedings of the National Academy of Sciences of the United*

States of America 102 (14): 4936–42.
Li, Bo, Sei Kuriyama, Mauricio Moreno, and Roberto Mayor. 2009. “The Posteriorizing Gene Gbx2 Is a Direct Target of Wnt Signalling and the Earliest Factor in Neural Crest Induction.” *Development (Cambridge, England)* 136 (19). The Company of Biologists Ltd: 3267–78.
Li, Gene-Wei, David Burkhardt, Carol Gross, and Jonathan S Weissman. 2014. “Quantifying Absolute Protein Synthesis Rates Reveals Principles Underlying Allocation of Cellular Resources.” *Cell* 157 (3). NIH Public Access: 624–35.
Li, N., K. Nakamura, Yi Jiang, Hiromichi Tsurui, Shuji Matsuoka, Masaaki Abe, Mareki Ohtsuji, et al. 2003. “Gain-of-Function Polymorphism in Mouse and Human Ltk: Implications for the Pathogenesis of Systemic Lupus Erythematosus.” *Human Molecular Genetics* 13 (2). Oxford University Press: 171–79.
Li, Wei, and Robert A. Cornell. 2007. “Redundant Activities of Tfap2a and Tfap2c Are Required for Neural Crest Induction and Development of Other Non-Neural Ectoderm Derivatives in Zebrafish Embryos.” *Developmental Biology* 304 (1): 338–54.
Li, Yin-Xiong, Marzena Zdanowicz, Lori Young, Donna Kumiski, Linda Leatherbury, and Margaret L Kirby. 2003. “Cardiac Neural Crest in Zebrafish Embryos Contributes to Myocardial Cell Lineage and Early Heart Function.” *Developmental Dynamics : An Official Publication of the American Association of Anatomists* 226 (3): 540–50.
Liem, K F, G Tremml, H Roelink, and T M Jessell. 1995. “Dorsal Differentiation of Neural Plate Cells Induced by BMP-Mediated Signals from Epidermal Ectoderm.” *Cell* 82 (6): 969–79.
Ling, F., B. Kang, and X Sun. 2014. “Id Proteins: Small Molecules, Mighty Regulators.” In *bHLH Transcription Factors in Development and Disease*, edited by R. Taneja, 1st ed., 398. Elsevier Science.
Lister, J A, J Close, and D W Raible. 2001. “Duplicate Mitf Genes in Zebrafish: Complementary Expression and Conservation of Melanogenic Potential.” *Developmental Biology* 237 (2): 333–44.
Lister, J A, C P Robertson, T Lepage, S L Johnson, and D W Raible. 1999. “Nacre Encodes a Zebrafish Microphthalmia-Related Protein That Regulates Neural-Crest-Derived Pigment Cell Fate.” *Development* 126 (17): 3757–67.
Lister, James A, Cynthia Cooper, Kim Nguyen, Melinda Modrell, Kelly Grant, and David W Raible. 2006. “Zebrafish Foxd3 Is Required for Development of a Subset of Neural Crest Derivatives.” *Developmental Biology* 159 (1). Elsevier Inc.: 50–59.
Lister, James A, Brandon M Lane, Anhthu Nguyen, and Katherine Lunney. 2011. “Embryonic Expression of Zebrafish MiT Family Genes tfe3b, Tfeb, and Tfec.” *Developmental Dynamics* 240 (11): 2529–38.
Liu, Yi-Wen. 2007. “Interrenal Organogenesis in the Zebrafish Model.” *Organogenesis* 3 (1). Landes Bioscience: 44–48.
Liu, Ying, and Patricia A. Labosky. 2008. “Regulation of Embryonic Stem Cell Self-Renewal and Pluripotency by Foxd3.” *Stem Cells* 26 (10). John Wiley & Sons, Ltd.: 2475–84.
Lo, Liching, Emma Dormand, Amy Greenwood, and David J Anderson. 2002. “Comparison of the Generic Neuronal Differentiation and Neuron Subtype Specification Functions of Mammalian Achaete-Scute and Atonal Homologs in Cultured Neural Progenitor Cells.” *Development (Cambridge, England)* 129 (7): 1553–67.
Locke, James C W, Jonathan W Young, Michelle Fontes, María Jesús Hernández Jiménez, and Michael B Elowitz. 2011. “Stochastic Pulse Regulation in Bacterial Stress Response.” *Science (New York, N.Y.)* 334 (6054). NIH Public Access: 366–69.
Lopes, Susana S, Xueyan Yang, Jeanette Müller, Thomas J Carney, Anthony R McAdow, Gerd-Jörg Rauch, Arie S Jacoby, et al. 2008. “Leukocyte Tyrosine Kinase Functions in Pigment Cell Development.” Edited by Greg Barsh. *PLoS Genetics* 4 (3). Public Library of Science: 13.

Ludwig, Andreas, Stephan Rehberg, and Michael Wegner. 2004. “Melanocyte-Specific Expression of Dopachrome Tautomerase Is Dependent on Synergistic Gene Activation by the Sox10 and Mitf Transcription Factors.” *FEBS Letters* 556 (1–3): 236–44.
Lumsden, A, N Sprawson, and A Graham. 1991. “Segmental Origin and Migration of Neural Crest Cells in the Hindbrain Region of the Chick Embryo.” *Development (Cambridge, England)* 113 (4). The Company of Biologists Ltd: 1281–91.
Ma, Long, Lei Lei, S Raisa Eng, Eric Turner, and Luis F Parada. 2003. “Brn3a Regulation of TrkA/NGF Receptor Expression in Developing Sensory Neurons.” *Development (Cambridge, England)* 130 (15): 3525–34.
Ma, Q, C Kintner, and D J Anderson. 1996. “Identification of Neurogenin, a Vertebrate Neuronal Determination Gene.” *Cell* 87 (1): 43–52.
Mahony, Christopher B, Richard J Fish, Corentin Pasche, and Julien Y Bertrand. 2016. “Tfec Controls the Hematopoietic Stem Cell Vascular Niche during Zebrafish Embryogenesis.” *Blood* 128 (10). American Society of Hematology: 1336–45.
Mallamaci, Antonello, Erica Di Blas, Paola Briata, Edoardo Boncinelli, and Giorgio Corte. 1996. “OTX2 Homeoprotein in the Developing Central Nervous System and Migratory Cells of the Olfactory Area.” *Mechanisms of Development* 58 (1): 165–78.
Marchal, L., G. Luxardi, V. Thome, and L. Kodjabachian. 2009. “BMP Inhibition Initiates Neural Induction via FGF Signaling and Zic Genes.” *Proceedings of the National Academy of Sciences* 106 (41). National Academy of Sciences: 17437–42.
Marcucio, Ralph S, Nathan M Young, Diane Hu, and Benedikt Hallgrimsson. 2011. “Mechanisms That Underlie Co-Variation of the Brain and Face.” *Genesis (New York, N.Y. : 2000)* 49 (4). NIH Public Access: 177–89.
Marie, Sandrine, Benedicte Heron, Pierre Bitoun, Therese Timmerman, Georges Van Den Berghe, and Marie-Francoise Vincent. 2004. “AICA-Ribosiduria: A Novel, Neurologically Devastating Inborn Error of Purine Biosynthesis Caused by Mutation of ATIC.” *American Journal of Human Genetics* 74 (6). Elsevier: 1276–81.
Martina, José A, Heba I Diab, Huiqing Li, and Rosa Puertollano. 2014. “Novel Roles for the MiTF/TFE Family of Transcription Factors in Organelle Biogenesis, Nutrient Sensing, and Energy Homeostasis.” *Cellular and Molecular Life Sciences : CMLS* 71 (13): 2483–97.
Maru, Y, H Hirai, and F Takaku. 1990. “Human Ltk: Gene Structure and Preferential Expression in Human Leukemic Cells.” *Oncogene Research* 5 (3): 199–204.
Matsumoto, Jiro. 1965. “Studies on Fine Structure and Cytochemical Properties of Erythrophores in Swordtail, Xiphophorus Helleri, with Special Reference to Their Pigment Granules (Pterinosomes).” *The Journal of Cell Biology* 27 (3). The Rockefeller University Press: 493–504.
Mayor, R, N Guerrero, and C Martínez. 1997. “Role of FGF and Noggin in Neural Crest Induction.” *Developmental Biology* 189 (1): 1–12.
McCallion, A S, and A Chakravarti. 2001. “EDNRB/EDN3 and Hirschsprung Disease Type II.” *Pigment Cell Research / Sponsored by the European Society for Pigment Cell Research and the International Pigment Cell Society* 14 (3): 161–69.
McGraw, Hillary F, Alexei Nechiporuk, and David W Raible. 2008. “Zebrafish Dorsal Root Ganglia Neural Precursor Cells Adopt a Glial Fate in the Absence of neurogenin1.” *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience* 28 (47): 12558–69.
McLennan, Rebecca, Linus J Schumacher, Jason A Morrison, Jessica M Teddy, Dennis A Ridenour, Andrew C Box, Craig L Semerad, et al. 2015. “Neural Crest Migration Is Driven by a Few Trailblazer Cells with a Unique Molecular Signature Narrowly Confined to the Invasive Front.” *Development (Cambridge, England)* 142 (11). Oxford University Press for The Company of Biologists Limited: 2014–25.
Menter, David G, M Obika, T T Tchen, and John D Taylor. 1979. “Leucophores and Iridophores of Fundulus Heteroclitus : Biophysical and Ultrastructural Properties.” *JMorphol* 160 (1). Wiley Subscription Services, Inc., A Wiley Company: 120.

Mertin, Sabine, Sharon G. McDowall, and Vincent R. Harley. 1999. “The DNA-Binding Specificity of SOX9 and Other SOX Proteins.” *Nucleic Acids Research* 27 (5): 1359–64.
Minoux, Maryline, and Filippo M Rijli. 2010. “Molecular Mechanisms of Cranial Neural Crest Cell Migration and Patterning in Craniofacial Development.” *Development (Cambridge, England)* 137 (16): 2605–21.
Mongera, Alessandro, Ajeet P Singh, Mitchell P Levesque, Yi-Yen Chen, Peter Konstantinidis, and Christiane Nüsslein-Volhard. 2013. “Genetic Lineage Labeling in Zebrafish Uncovers Novel Neural Crest Contributions to the Head, Including Gill Pillar Cells.” *Development (Cambridge, England)* 140 (4): 916–25.
Monsoro-Burq, Anne-Hélène, Estee Wang, and Richard Harland. 2005. “Msx1 and Pax3 Cooperate to Mediate FGF8 and WNT Signals during Xenopus Neural Crest Induction.” *Developmental Cell* 8 (2): 167–78.
Montero-Balaguer, Mercedes, Michael R Lang, Sherri Weiss Sachdev, Christiane Knappmeyer, Rodney A Stewart, Ana De La Guardia, Antonis K Hatzopoulos, and Ela W Knapik. 2006. “The Mother Superior Mutation Ablates foxd3 Activity in Neural Crest Progenitor Cells and Depletes Neural Crest Derivatives in Zebrafish.” *Developmental Dynamics an Official Publication of the American Association of Anatomists* 235 (12): 3199–3212.
Montoliu, Lluís, Karen Grønskov, Ai-Hua Wei, Mónica Martínez-García, Almudena Fernández, Benoît Arveiler, Fanny Morice-Picard, et al. 2014. “Increasing the Complexity: New Genes and New Types of Albinism.” *Pigment Cell & Melanoma Research* 27 (1): 11–18.
Morin-Kensicki, E M, and J S Eisen. 1997. “Sclerotome Development and Peripheral Nervous System Segmentation in Embryonic Zebrafish.” *Development (Cambridge, England)* 124 (1): 159–67.
Moriwaki, Yuji, T. Yamamoto, and K. Higashino. 1999. “Enzymes Involved in Purine Metabolism - A Review of Histochemical Localization and Functional Implications.” *Histology and Histopathology* 14 (4): 1321–40.
Morriss-Kay, G, E Ruberte, and Y Fukiishi. 1993. “Mammalian Neural Crest and Neural Crest Derivatives.” *Annals Of Anatomy* 175 (6): 501–7.
Mort, Richard L, Ian J Jackson, and E Elizabeth Patton. 2015. “The Melanocyte Lineage in Development and Disease.” *Development (Cambridge, England)* 142 (4). Company of Biologists: 620–32.
Mundell, Nathan A, and Patricia A Labosky. 2011. “Neural Crest Stem Cell Multipotency Requires Foxd3 to Maintain Neural Potential and Repress Mesenchymal Fates.” *Development* 138 (4): 641–52.
Murugan, Rajamanickam. 2012. “Theory on the Dynamics of Feedforward Loops in the Transcription Factor Networks.” Edited by Christof Markus Aegerter. *PLoS One* 7 (7). Public Library of Science: e41027.
Nadig, Roopa R. 2009. “Stem Cell Therapy - Hype or Hope? A Review.” *Journal of Conservative Dentistry* 12 (4): 131–38.
Nagatomo, Kan-Ichiro, and Chikara Hashimoto. 2007. “Xenopus hairy2 Functions in Neural Crest Formation by Maintaining Cells in a Mitotic and Undifferentiated State.” *Developmental Dynamics : An Official Publication of the American Association of Anatomists* 236 (6): 1475–83.
Nasevicius, A, and S C Ekker. 2000. “Effective Targeted Gene ‘Knockdown’ in Zebrafish.” *Nature Genetics* 26 (2): 216–20.
Ng, Anthony, Rosa A Uribe, Leah Yieh, Richard Nuckels, and Jeffrey M Gross. 2009. “Zebrafish Mutations in Gart and Paics Identify Crucial Roles for de Novo Purine Synthesis in Vertebrate Pigmentation and Ocular Development.” *Development (Cambridge, England)* 136 (15). Company of Biologists: 2601–11.
Nichane, Massimo, Noémie de Crozé, Xi Ren, Jacob Souopgui, Anne H Monsoro-Burq, and Eric J Bellefroid. 2008. “Hairy2-Id3 Interactions Play an Essential Role in Xenopus Neural Crest Progenitor Specification.” *Developmental Biology* 322 (2): 355–67.
O’Brien, Erin K, Claudia D’Alençon, Gregory Bonde, Wei Li, Jeff Schoenebeck, Miguel

L Allende, Bruce D Gelb, Deborah Yelon, Judith S Eisen, and Robert A Cornell. 2004. “Transcription Factor Ap-2alpha Is Necessary for Development of Embryonic Melanophores, Autonomic Neurons and Pharyngeal Skeleton in Zebrafish.” *Developmental Biology* 265 (1): 246–61.

Obika, Masataka. 1993. “Formation of Pterinosomes and Carotenoid Granules in Xanthophores of the Teleost Oryzias Latipes as Revealed by the Rapid-Freezing and Freeze-Substitution Method.” *Cell & Tissue Research* 271 (1): 81–86.

Odenthal, J, P Haffter, E Vogelsang, M Brand, F J Van Eeden, M Furutani-Seiki, M Granato, et al. 1996. “Mutations Affecting the Formation of the Notochord in the Zebrafish, Danio Rerio.” *Development* 123 (Sp. Iss. SI): 103–15.

Odenthal, J, and C Nüsslein-Volhard. 1998. “Fork Head Domain Genes in Zebrafish.” *Development Genes and Evolution* 208 (5): 245–58.

Odenthal, J, K Rossnagel, P Haffter, R N Kelsh, E Vogelsang, M Brand, F J Van Eeden, et al. 1996. “Mutations Affecting Xanthophore Pigmentation in the Zebrafish, Danio Rerio.” *Development* 123 (Sp. Iss. SI): 391–98.

Oetting, W S. 2000. “The Tyrosinase Gene and Oculocutaneous Albinism Type 1 (OCA1): A Model for Understanding the Molecular Biology of Melanin Formation.” *Pigment Cell Research / Sponsored by the European Society for Pigment Cell Research and the International Pigment Cell Society* 13 (5): 320–25.

Olden, Tasha, Tanveer Akhtar, Sarah A. Beckman, and Kenneth N. Wallace. 2008. “Differentiation of the Zebrafish Enteric Nervous System and Intestinal Smooth Muscle.” *Genesis* 46 (9). Wiley Subscription Services, Inc., A Wiley Company: 484–98.

Orkin, Stuart H., and Leonard I. Zon. 2008. “Hematopoiesis: An Evolving Paradigm for Stem Cell Biology.” *Cell* 132 (4): 631–44.

Otterloo, Eric Van, Wei Li, Aaron Garnett, Maria Cattell, Daniel Meulemans Medeiros, and Robert A Cornell. 2012. “Novel Tfap2-Mediated Control of soxE Expression Facilitated the Evolutionary Emergence of the Neural Crest.” *Development* 139 (4): 720–30.

Parichy, D M, E M Mellgren, J F Rawls, S S Lopes, R N Kelsh, and S L Johnson. 2010. “Mutational Analysis of Endothelin Receptor b1 (Rose) during Neural Crest and Pigment Pattern Development in the Zebrafish Danio Rerio.” *Developmental Biology* 227 (2): 294–306.

Pavan, William J, and David W Raible. 2012. “Specification of Neural Crest into Sensory Neuron and Melanocyte Lineages.” *Developmental Biology* 366 (1): 55–63.

Pearson, Karl. 1900. “X. On the Criterion That a given System of Deviations from the Probable in the Case of a Correlated System of Variables Is Such That It Can Be Reasonably Supposed to Have Arisen from Random Sampling.” *Philosophical Magazine Series 5* 50 (302). Taylor & Francis: 157–75.

Peirano, Reto I, and Michael Wegner. 2000. “The Glial Transcription Factor Sox10 Binds to DNA Both as Monomer and Dimer with Different Functional Consequences.” *Nucleic Acids Research* 28 (16). Oxford University Press: 3047–55.

Peng, X, M Dong, L Ma, X-E Jia, J Mao, C Jin, Y Chen, et al. 2015. “A Point Mutation of Zebrafish c-Cbl Gene in the Ring Finger Domain Produces a Phenotype Mimicking Human Myeloproliferative Disease.” *Leukemia* 29 (12). Nature Publishing Group: 2355–65.

Pera, E, S Stein, and M Kessel. 1999. “Ectodermal Patterning in the Avian Embryo: Epidermis versus Neural Plate.” *Development (Cambridge, England)* 126 (1): 63–73.

Plickert, G, M Kroiher, and A Munck. 1988. “Cell Proliferation and Early Differentiation during Embryonic Development and Metamorphosis of Hydractinia Echinata.” *Development (Cambridge, England)* 103 (4): 795–803.

Ploper, Diego, and Edward M. De Robertis. 2015. “The MITF Family of Transcription Factors: Role in Endolysosomal Biogenesis, Wnt Signaling, and Oncogenesis.” *Pharmacological Research* 99: 36–43.

Plouhinec, Jean-Louis, Daniel D. Roche, Caterina Pegoraro, Ana Leonor Figueiredo, Frédérique Maczkowiak, Lisa J. Brunet, Cécile Milet, et al. 2014. “Pax3 and Zic1 Trigger the Early Neural Crest Gene Regulatory Network by the Direct Activation of Multiple Key Neural Crest Specifiers.” *Developmental Biology* 386 (2): 461–72.
Pogenberg, Vivian, Margrét H. Ögmundsdóttir, Kristín Bergsteinsdóttir, Alexander Schepsky, Bengt Phung, Viktor Deineko, Morlin Milewski, Eiríkur Steingrímsson, and Matthias Wilmanns. 2012. “Restricted Leucine Zipper Dimerization and Specificity of DNA Recognition of the Melanocyte Master Regulator MITF.” *Genes and Development* 26 (23): 2647–58.
Pohl, B S, and W Knöchel. 2001. “Overexpression of the Transcriptional Repressor FoxD3 Prevents Neural Crest Formation in Xenopus Embryos.” *Mechanisms of Development* 103 (1–2): 93–106.
Potterf, S B, M Furumura, K J Dunn, H Arnheiter, and W J Pavan. 2000. “Transcription Factor Hierarchy in Waardenburg Syndrome: Regulation of MITF Expression by SOX10 and PAX3.” *Human Genetics* 107 (1): 1–6.
Powell, Davalyn R, Laura Hernandez-Lagunas, Kristi LaMonica, and Kristin Bruk Artinger. 2013. “Prdm1a Directly Activates foxd3 and tfap2a during Zebrafish Neural Crest Specification.” *Development (Cambridge, England)* 140 (16): 3445–55.
Prendergast, Andrew, and David W. Raible. 2014. “Chapter 13 – Neural Crest Cells and Peripheral Nervous System Development.” In *Neural Crest Cells*, 255–86.
Purves, Dale, George J. Augustine, David Fitzpatrick, William C. Hall, Anthony-Samuel LaMantia, and Leonard E. White. 2012. *Neuroscience*. 5th ed. Sinauer Associates, Inc.
Quigley, Ian K, and David M Parichy. 2002. “Pigment Pattern Formation in Zebrafish: A Model for Developmental Genetics and the Evolution of Form.” *Microscopy Research and Technique* 58 (6). New York, NY: Wiley-Liss, c1992-: 442–55.
Rada-Iglesias, Alvaro, Ruchi Bajpai, Sara Prescott, Samantha A. Brugmann, Tomek Swigut, Joanna Wysocka, J. Amiel, et al. 2012. “Epigenomic Annotation of Enhancers Predicts Transcriptional Regulators of Human Neural Crest.” *Cell Stem Cell* 11 (5). Elsevier: 633–48.
Raible, David W., Andrew Wood, Wendy Hodsdon, Paul D. Henion, James A. Weston, and Judith S. Eisen. 1992. “Segregation and Early Dispersal of Neural Crest Cells in the Embryonic Zebrafish.” *Developmental Dynamics* 195 (1). Wiley Subscription Services, Inc., A Wiley Company: 29–42.
Raible, David W, and Josette M Ungos. 2006. “Specification of Sensory Neuron Cell Fate from the Neural Crest.” *Advances in Experimental Medicine and Biology* 589 (January): 170–80.
Ramgolam, Kiran, Jessica Lauriol, Claude Lalou, Laura Lauden, Laurence Michel, Pierre de la Grange, Abdel-Majid Khatib, et al. 2011. “Melanoma Spheroids Grown Under Neural Crest Cell Conditions Are Highly Plastic Migratory/Invasive Tumor Cells Endowed with Immunomodulator Function.” Edited by Benjamin Edward Rich. *PLoS ONE* 6 (4). Public Library of Science: e18784.
Rau, Marlene J., Sabine Fischer, and Carl J. Neumann. 2006. “Zebrafish Trap230/Med12 Is Required as a Coactivator for Sox9-Dependent Neural Crest, Cartilage and Ear Development.” *Developmental Biology* 296 (1): 83–93.
Rauch, Gerd-Jörg, Michael Granato, and Pascal Haffter. 1997. “A Polymorphic Zebrafish Line for Genetic Mapping Using SSLPs on High-Percentage Agarose Gels.” *Technical Tips Online* 2 (1): 148–50.
Rawls, John F, and Stephen L Johnson. 2003. “Temporal and Molecular Separation of the Kit Receptor Tyrosine Kinase’s Roles in Zebrafish Melanocyte Migration and Survival.” *Developmental Biology* 262 (1): 152–61.
Read, A P, and V E Newton. 1997. “Waardenburg Syndrome.” *Journal of Medical Genetics* 34 (8). BMJ Publishing Group Ltd: 656–65.
Rehli, M, A Lichanska, A I Cassady, M C Ostrowski, and D A Hume. 1999. “TFEC Is a Macrophage-Restricted Member of the Microphthalmia-TFE Subfamily of Basic Helix-Loop-Helix Leucine Zipper Transcription Factors.” *Journal of Immunology*

(Baltimore, Md. : 1950) 162 (3): 1559–65.
Reshetnyak, Andrey V, Phillip B Murray, Xiarong Shi, Elizabeth S Mo, Jyotidarsini Mohanty, Francisco Tome, Hanwen Bai, Murat Gunel, Irit Lax, and Joseph Schlessinger. 2015. “Augmentor α and β (FAM150) Are Ligands of the Receptor Tyrosine Kinases ALK and LTK: Hierarchy and Specificity of Ligand-Receptor Interactions.” *Proceedings of the National Academy of Sciences of the United States of America* 112 (52). National Academy of Sciences: 15862–67.
Respuela, Patricia, Miloš Nikolić, Minjia Tan, Peter Frommolt, Yingming Zhao, Joanna Wysocka, and Alvaro Rada-Iglesias. 2016. “Foxd3 Promotes Exit from Naive Pluripotency through Enhancer Decommissioning and Inhibits Germline Specification.” *Cell Stem Cell* 18 (1): 118–33.
Rex, M., A. Orme, Dafe Uwanogho, Kevin Tointon, Peter M. Wigmore, Paul T. Sharpe, and Paul J. Scotting. 1997. “Dynamic Expression of Chicken Sox2 and Sox3 Genes in Ectoderm Induced to Form Neural Tissue.” *Developmental Dynamics* 209 (3). Wiley- Liss, Inc.: 323–32.
Ribes, Vanessa, and James Briscoe. 2009. “Establishing and Interpreting Graded Sonic Hedgehog Signaling during Vertebrate Neural Tube Patterning: The Role of Negative Feedback.” *Cold Spring Harbor Perspectives in Biology* 1 (2). Cold Spring Harbor Laboratory Press: a002014.
Riley, P A. 1999. “The Great DOPA Mystery: The Source and Significance of DOPA in Phase I Melanogenesis.” *Cellular and Molecular Biology (Noisy-Le-Grand, France)* 45 (7): 951–60.
Rinon, Ariel, Shlomi Lazar, Heather Marshall, Stine Büchmann-Møller, Adi Neufeld, Hadas Elhanany-Tamir, Makoto M Taketo, Lukas Sommer, Robb Krumlauf, and Eldad Tzahor. 2007. “Cranial Neural Crest Cells Regulate Head Muscle Patterning and Differentiation during Vertebrate Embryogenesis.” *Development (Cambridge, England)* 134 (17): 3065–75.
Robu, Mara E, Jon D Larson, Aidas Nasevicius, Soraya Beiraghi, Charles Brenner, Steven A Farber, and Stephen C Ekker. 2007. “p53 Activation by Knockdown Technologies.” *PLoS Genetics* 3 (5): e78.
Rodrigues, Frederico S, Gail Doughton, B Yang, and Robert N Kelsh. 2012. “A Novel Transgenic Line Using the Cre-Lox System to Allow Permanent Lineage-Labeling of the Zebrafish Neural Crest.” *Genesis* 8: 1–8.
Roskoski, Robert. 2007. “Sunitinib: A VEGF and PDGF Receptor Protein Kinase and Angiogenesis Inhibitor.” *Biochemical and Biophysical Research Communications* 356 (2): 323–28.
Rossi, Andrea, Zacharias Kontarakis, Claudia Gerri, Hendrik Nolte, Soraya Hölper, Marcus Krüger, and Didier Y R Stainier. 2015. “Genetic Compensation Induced by Deleterious Mutations but Not Gene Knockdowns.” *Nature* 524 (7564): 230–33.
Saito, Daisuke, Yuta Takase, Hidetaka Murai, Yoshiko Takahashi, N. Le Douarin, M. A. Teillet, M. A. Teillet, et al. 2012. “The Dorsal Aorta Initiates a Molecular Cascade That Instructs Sympatho-Adrenal Specification.” *Science (New York, N.Y.)* 336 (6088). American Association for the Advancement of Science: 1578–81.
Saldana-Caboverde, Amy, and Lidia Kos. 2010. “Roles of Endothelin Signaling in Melanocyte Development and Melanoma.” *Pigment Cell & Melanoma Research* 23 (2). NIH Public Access: 160–70.
Sander, Jeffry D, and J Keith Joung. 2014. “CRISPR-Cas Systems for Editing, Regulating and Targeting Genomes.” *Nature Biotechnology* 32 (4). Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.: 347–55.
Santagati, Fabio, and Filippo M Rijli. 2003. “Cranial Neural Crest and the Building of the Vertebrate Head.” *Nature Reviews. Neuroscience* 4 (10): 806–18.
Sasai, N, K Mizuseki, and Y Sasai. 2001. “Requirement of FoxD3-Class Signaling for Neural Crest Determination in Xenopus.” *Development* 128 (13): 2525–36.
Sato, Takahiko, Noriaki Sasai, and Yoshiki Sasai. 2005. “Neural Crest Determination by Co-Activation of Pax3 and Zic1 Genes in Xenopus Ectoderm.” *Development (Cambridge, England)* 132 (10). The Company of Biologists Ltd: 2355–63.

Sauka-Spengler, T., and M. Bronner-Fraser. 2008. “A Gene Regulatory Network Orchestrates Neural Crest Formation.” *Nature Reviews Molecular Cell Biology* 9 (7): 557–68.
Schartl, Manfred, Lionel Larue, Makoto Goda, Marcus W. Bosenberg, Hisashi Hashimoto, and Robert N. Kelsh. 2016. “What Is a Vertebrate Pigment Cell?” *Pigment Cell & Melanoma Research* 29 (1): 8–14.
Schilling, T F, and C B Kimmel. 1994. “Segment and Cell Type Lineage Restrictions during Pharyngeal Arch Development in the Zebrafish Embryo.” *Development (Cambridge, England)* 120 (3). The Company of Biologists Ltd: 483–94.
Schlierf, Beate, Andreas Ludwig, Karin Klenovsek, and Michael Wegner. 2002. “Cooperative Binding of Sox10 to DNA: Requirements and Consequences.” *Nucleic Acids Research* 30 (24): 5509–16.
Schulte-Merker, S., and D. Y. R. Stainier. 2014. “Out with the Old, in with the New: Reassessing Morpholino Knockdowns in Light of Genome Editing Technology.” *Development* 141 (16): 3103–4.
Schumacher, Jennifer A, Megumi Hashiguchi, Vu H Nguyen, and Mary C Mullins. 2011. “An Intermediate Level of BMP Signaling Directly Specifies Cranial Neural Crest Progenitor Cells in Zebrafish.” *PloS One* 6 (11). Public Library of Science: e27403.
Schwanhäusser, Björn, Dorothea Busse, Na Li, Gunnar Dittmar, Johannes Schuchhardt, Jana Wolf, Wei Chen, and Matthias Selbach. 2011. “Global Quantification of Mammalian Gene Expression Control.” *Nature* 473 (7347): 337–42.
Sérandour, Aurélien A, Stéphane Avner, Frédéric Percevault, Florence Demay, Maud Bizot, Céline Lucchetti-Miganeh, Frédérique Barloy-Hubler, et al. 2011. “Epigenetic Switch Involved in Activation of Pioneer Factor FOXA1-Dependent Enhancers.” *Genome Research* 21 (4): 555–65.
Serbedzija, G N, S E Fraser, and M Bronner-Fraser. 1990. “Pathways of Trunk Neural Crest Cell Migration in the Mouse Embryo as Revealed by Vital Dye Labelling.” *Development (Cambridge, England)* 108 (4): 605–12.
Shakhova, Olga, Phil Cheng, Pravin J Mishra, Daniel Zingg, Simon M Schaefer, Julien Debbache, Jessica Häusel, et al. 2015. “Antagonistic Cross-Regulation between Sox9 and Sox10 Controls an Anti-Tumorigenic Program in Melanoma.” *PLoS Genetics* 11 (1): e1004877.
Shakhova, Olga, and Lukas Sommer. 2008. *Neural Crest-Derived Stem Cells*. *StemBook*. Harvard Stem Cell Institute.
Shamblott, Michael J., Marci L. O’Driscoll, Danielle L. Gomez, Dustin L. McGuire, JC Lee, SB Smith, H Watada, et al. 2016. “Neurogenin 3 Is Regulated by Neurotrophic Tyrosine Kinase Receptor Type 2 (TRKB) Signaling in the Adult Human Exocrine Pancreas.” *Cell Communication and Signaling* 14 (1). BioMed Central: 23.
Sheng, G, and C D Stern. 1999. “Gata2 and Gata3: Novel Markers for Early Embryonic Polarity and for Non-Neural Ectoderm in the Chick Embryo.” *Mechanisms of Development* 87 (1–2): 213–16.
Shepherd, Iain, and Judith Eisen. 2011. “Development of the Zebrafish Enteric Nervous System.” *Methods in Cell Biology* 101. NIH Public Access: 143–60.
Simoes-Costa, M., J. Tan-Cabugao, I. Antoshechkin, T. Sauka-Spengler, and M. E. Bronner. 2014. “Transcriptome Analysis Reveals Novel Players in the Cranial Neural Crest Gene Regulatory Network.” *Genome Research* 24 (2). Cold Spring Harbor Laboratory Press: 281–90.
Simões-Costa, Marcos, Marianne E Bronner, P. Agarwal, M. P. Verzi, T. Nguyen, J. Hu, M. L. Ehlers, et al. 2015. “Establishing Neural Crest Identity: A Gene Regulatory Recipe.” *Development (Cambridge, England)* 142 (2). Oxford University Press for The Company of Biologists Limited: 242–57.
Simões-Costa, Marcos S., Sonja J. McKeown, Joanne Tan-Cabugao, Tatjana Sauka-Spengler, Marianne E. Bronner, T Sauka-Spengler, M Barembaum, et al. 2012. “Dynamic and Differential Regulation of Stem Cell Factor FoxD3 in the Neural

Crest Is Encrypted in the Genome." Edited by Hisato Kondoh. *PLoS Genetics* 8 (12). Public Library of Science: e1003142.
Smith, S. D, Philip M Kelley, Judith B Kenyon, and Denise Hoover. 2000. "Tietz Syndrome (Hypopigmentation/deafness) Caused by Mutation of MITF." *Journal of Medical Genetics* 37 (6). BMJ Publishing Group Ltd: 446–48.
Spitz, François, and Eileen E. M. Furlong. 2012. "Transcription Factors: From Enhancer Binding to Developmental Control." *Nature Reviews Genetics* 13 (9). Nature Publishing Group: 613–26.
Stemple, D L, and D J Anderson. 1992. "Isolation of a Stem Cell for Neurons and Glia from the Mammalian Neural Crest." *Cell* 71 (6): 973–85.
Steventon, Ben, Claudio Araya, Claudia Linker, Sei Kuriyama, and Roberto Mayor. 2009. "Differential Requirements of BMP and Wnt Signalling during Gastrulation and Neurulation Define Two Steps in Neural Crest Induction." *Development (Cambridge, England)* 136 (5): 771–79.
Steventon, Ben, and Roberto Mayor. 2012. "Early Neural Crest Induction Requires an Initial Inhibition of Wnt Signals." *Developmental Biology* 365 (1): 196–207.
Steventon, Ben, Roberto Mayor, and Andrea Streit. 2014. "Neural Crest and Placode Interaction during the Development of the Cranial Sensory System." *Developmental Biology* 389 (1). NIH Public Access: 28–38.
Stewart, Rodney A, Brigitte L Arduini, Stephane Berghmans, Rani E George, John P Kanki, Paul D Henion, and A Thomas Look. 2006. "Zebrafish foxd3 Is Selectively Required for Neural Crest Specification, Migration and Survival." *Developmental Biology* 292 (1). Elsevier Inc.: 174–88.
Stewart, Rodney A, Jeong-Soo Lee, Martina Lachnit, A Thomas Look, John P Kanki, and Paul D Henion. 2010. "Studying Peripheral Sympathetic Nervous System Development and Neuroblastoma in Zebrafish." *Methods in Cell Biology* 100 (January): 127–52.
Stoller, Jason Z, and Jonathan A Epstein. 2005. "Cardiac Neural Crest." *Seminars in Cell & Developmental Biology* 16 (6): 704–15.
Streit, A, and C D Stern. 1999. "Establishment and Maintenance of the Border of the Neural Plate in the Chick: Involvement of FGF and BMP Activity." *Mechanisms of Development* 82 (1–2): 51–66.
Streit, Andrea, Alyson J. Berliner, Costis Papanayotou, Andrés Sirulnik, and Claudio D. Stern. 2000. "Initiation of Neural Induction by FGF Signalling before Gastrulation." *Nature* 406 (6791). Nature Publishing Group: 74–78.
Suzuki, A, N Ueno, and A Hemmati-Brivanlou. 1997. "Xenopus msx1 Mediates Epidermal Induction and Neural Inhibition by BMP4." *Development (Cambridge, England)* 124 (16): 3037–44.
Sweet, Deborah J., T. Boroviak, R. Loos, P. Lombard, J. Okahara, R. Behr, E. Sasaki, et al. 2016. "Foxd3: A Repressor, an Activator, or Both?" *Cell Stem Cell* 18 (1). Elsevier: 1–2.
Swiers, Gemma, Roger Patient, and Matthew Loose. 2006. "Genetic Regulatory Networks Programming Hematopoietic Stem Cells and Erythroid Lineage Specification." *Developmental Biology* 294 (2): 525–40.
Takada, Norio, and Bruce Appel. 2011. "swap70 Promotes Neural Precursor Cell Cycle Exit and Oligodendrocyte Formation." *Mol. Cell Neuroscience* 28 (3): 225–35.
Takemoto, Tatsuya, Masanori Uchikawa, Yusuke Kamachi, and Hisato Kondoh. 2006. "Convergence of Wnt and FGF Signals in the Genesis of Posterior Neural Plate through Activation of the Sox2 Enhancer N-1." *Development (Cambridge, England)* 133 (2): 297–306.
Talbot, Jared C, and Sharon L Amacher. 2014. "Supplementary Methods: A Streamlined CRISPR Pipeline to Reliably Generate Zebrafish Frameshifting Alleles." *Zebrafish* 11 (6).
Taneyhill, Lisa A., and Andrew T. Schiffmacher. 2013. "Cadherin Dynamics During Neural Crest Cell Ontogeny." In *Progress in Molecular Biology and Translational Science*, 116:291–315.
Teng, Lu, Nathan A Mundell, Audrey Y Frist, Qiaohong Wang, and Patricia A Labosky.

2008. “Requirement for Foxd3 in the Maintenance of Neural Crest Progenitors.” *Development (Cambridge, England)* 135 (9): 1615–24.
Theveneau, Eric, and Roberto Mayor. 2012. “Neural Crest Delamination and Migration: From Epithelium-to-Mesenchyme Transition to Collective Cell Migration.” *Developmental Biology* 366 (1): 34–54.
Thien, Christine B F, and Wallace Y Langdon. 2005. “Negative Regulation of PTK Signalling by Cbl Proteins.” *Growth Factors (Chur, Switzerland)* 23 (2): 161–67.
Thisse, B., S. Pflumio, M. Fürthauer, B. Loppin, V. Heyer, A. Degrave, R. Woehl, et al. 2001. “Expression of the Zebrafish Genome during Embryogenesis.” *ZFIN Direct Data Submission (Http://zfin.org).*
Thomas, Aaron J, and Carol A Erickson. 2009. “FOXD3 Regulates the Lineage Switch between Neural Crest-Derived Glial Cells and Pigment Cells by Repressing MITF through a Non-Canonical Mechanism.” *Development* 136 (11). Company of Biologists: 1849–58.
Tian, Tianhai, and Kate Smith-Miles. 2014. “Mathematical Modeling of GATA-Switching for Regulating the Differentiation of Hematopoietic Stem Cell.” *BMC Systems Biology* 8 Suppl 1 (Suppl 1). BioMed Central: S8.
Tomlin, Claire J., and Jeffrey D. Axelrod. 2007. “Biology by Numbers: Mathematical Modelling in Developmental Biology.” *Nature Reviews Genetics* 8 (5). Nature Publishing Group: 331–40.
Trainor, Paul A. 2014. *Neural Crest Cells : Evolution, Development and Disease*. Elsevier/AP.
Tropepe, Vincent, Shuhong Li, Amanda Dickinson, Joshua T Gamse, and Hazel L Sive. 2006. “Identification of a BMP Inhibitor-Responsive Promoter Module Required for Expression of the Early Neural Gene zic1.” *Developmental Biology* 289 (2): 517–29.
Tuazon, Francesca B, and Mary C Mullins. 2015. “Temporally Coordinated Signals Progressively Pattern the Anteroposterior and Dorsoventral Body Axes.” *Seminars in Cell & Developmental Biology* 42 (June). NIH Public Access: 118–33.
Turing, A. M. 1952. “The Chemical Basis of Morphogenesis.” *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 237 (641).
Uribe, Rosa A, and Jeffrey M Gross. 2010. “Id2a Influences Neuron and Glia Formation in the Zebrafish Retina by Modulating Retinoblast Cell Cycle Kinetics.” *Development (Cambridge, England)* 137 (22): 3763–74.
Villaverde, A F, and J R Banga. 2013. “Reverse Engineering and Identification in Systems Biology: Strategies, Perspectives and Challenges.” *Journal of the Royal Society Interface*.
Vogel, K. S. 1996. “Origin of Adrenal Chromaffin Cells from the Neural Crest.” In *Genetic Mechanisms in Multiple Endocrine Neoplasia Type 2*, edited by D. Nelkin, 99–118. Springer Berlin Heidelberg.
Wang, Allen, Feng Yue, Yan Li, Ruiyu Xie, Thomas Harper, Nisha A Patel, Kayla Muth, et al. 2015. “Epigenetic Priming of Enhancers Predicts Developmental Competence of hESC-Derived Endodermal Lineage Intermediates.” *Cell Stem Cell* 16 (4). NIH Public Access: 386–99.
Wang, Fay, John Flanagan, Nan Su, Li-Chong Wang, Son Bui, Allissa Nielson, Xingyong Wu, Hong-Thuy Vo, Xiao-Jun Ma, and Yuling Luo. 2012. “RNAscope: A Novel in Situ RNA Analysis Platform for Formalin-Fixed, Paraffin-Embedded Tissues.” *The Journal of Molecular Diagnostics : JMD* 14 (1). American Society for Investigative Pathology: 22–29.
Wang, Xia, Alex K K Chan, Mai Har Sham, Alan J Burns, and Wood Yee Chan. 2011. “Analysis of the Sacral Neural Crest Cell Contribution to the Hindgut Enteric Nervous System in the Mouse Embryo.” *Gastroenterology* 141 (3): 992-1002-6.
Wang, Zhong, Mark Gerstein, and Michael Snyder. 2009. “RNA-Seq: A Revolutionary Tool for Transcriptomics.” *Nature Reviews. Genetics* 10 (1). NIH Public Access: 57–63.
Webb, Alexis B, Iván M Lengyel, David J Jörg, Guillaume Valentin, Frank Jülicher, Luis G Morelli, and Andrew C Oates. 2016. “Persistence, Period and Precision of

Autonomous Cellular Oscillators from the Zebrafish Segmentation Clock." *eLife* 5.
Weider, Matthias, and Michael Wegner. 2016. "SoxE Factors: Transcriptional Regulators of Neural Differentiation and Nervous System Development." *Seminars in Cell & Developmental Biology*, August.
Weiss, Joseph B, Changhui Xue, Ted Benice, Liquan Xue, Stephan W Morris, and Jacob Raber. 2012. "Anaplastic Lymphoma Kinase and Leukocyte Tyrosine Kinase: Functions and Genetic Interactions in Learning, Memory and Adult Neurogenesis." *Pharmacology, Biochemistry, and Behavior* 100 (3): 566–74.
Weiss, Michele B, Ethan V Abel, Neda Dadpey, and Andrew E Aplin. 2014. "FOXD3 Modulates Migration through Direct Transcriptional Repression of TWIST1 in Melanoma." *Molecular Cancer Research : MCR* 12 (9): 1314–23.
Westerfield, M. 2007. "The Zebrafish Book. A Guide for the Laboratory Use of Zebrafish (Danio Rerio), 5th Edition." *University of Oregon Press, Eugene (Book)*.
Weston, J A. 1991. "Sequential Segregation and Fate of Developmentally Restricted Intermediate Cell Populations in the Neural Crest Lineage." *Current Topics in Developmental Biology* 25 (January): 133–53.
Wetering, M van de, M Oosterwegel, K van Norren, and H Clevers. 1993. "Sox-4, an Sry-like HMG Box Protein, Is a Transcriptional Activator in Lymphocytes." *The EMBO Journal* 12 (10): 3847–54.
Wijchers, Patrick J. E. C., J. Peter H. Burbach, and Marten P. Smidt. 2006. "In Control of Biology: Of Mice, Men and Foxes." *Biochemical Journal* 397 (2).
Wilson, Sara, Anna Rydström, Tolleiv Trimborn, Karl Willert, Roel Nusse, Thomas M. Jessell, and Thomas Edlund. 2001. "The Status of Wnt Signalling Regulates Neural and Epidermal Fates in the Chick Embryo." *Nature* 411 (6835). Nature Publishing Group: 325–30.
Wittmann, Dominik M, Jan Krumsiek, Julio Saez-Rodriguez, Douglas A Lauffenburger, Steffen Klamt, Fabian J Theis, SA Kauffman, et al. 2009. "Transforming Boolean Models to Continuous Models: Methodology and Application to T-Cell Receptor Signaling." *BMC Systems Biology* 3 (1). BioMed Central: 98.
Woodhoo, Ashwin, and Lukas Sommer. 2008. "Development of the Schwann Cell Lineage: From the Neural Crest to the Myelinated Nerve." *Glia* 56 (14): 1481–90.
Yan, Yi-Lin, John Willoughby, Dong Liu, Justin Gage Crump, Catherine Wilson, Craig T Miller, Amy Singer, Charles Kimmel, Monte Westerfield, and John H Postlethwait. 2005. "A Pair of Sox: Distinct and Overlapping Functions of Zebrafish sox9 Co-Orthologs in Craniofacial and Pectoral Fin Development." *Development* 132 (5): 1069–83.
Yardley, Nathan, and Martín I García-Castro. 2012. "FGF Signaling Transforms Non-Neural Ectoderm into Neural Crest." *Developmental Biology* 372 (2): 166–77.
Zaret, Kenneth S, and Jason S Carroll. 2011. "Pioneer Transcription Factors: Establishing Competence for Gene Expression." *Genes & Development* 25 (21). Cold Spring Harbor Laboratory Press: 2227–41.
Zernicka-Goetz, Magdalena. 2002. "Patterning of the Embryo: The First Spatial Decisions in the Life of a Mouse." *Development (Cambridge, England)* 129 (4): 815–29.
Zhang, Hongbing, Lily I Pao, Aileen Zhou, Arthur D Brace, Robert Halenbeck, Amy W Hsu, Thomas L Bray, et al. 2014. "Deorphanization of the Human Leukocyte Tyrosine Kinase (LTK) Receptor by a Signaling Screen of the Extracellular Proteome." *Proceedings of the National Academy of Sciences of the United States of America* 111 (44): 15741–45.
Zhou, Qing, Hiram Chipperfield, Douglas A Melton, and Wing Hung Wong. 2007. "A Gene Regulatory Network in Mouse Embryonic Stem Cells." *Proceedings of the National Academy of Sciences of the United States of America* 104 (42): 16438–43.
Zhu, L., S. Zhang, Y. Jin, MJ. Evans, MH. Kaufman, I. Chambers, SR. Tomlinson, et al. 2014. "Foxd3 Suppresses NFAT-Mediated Differentiation to Maintain Self-Renewal of Embryonic Stem Cells." *EMBO Reports* 15 (12). EMBO Press: 1286–96.

Ziegler, Irmgard. 2003. “The Pteridine Pathway in Zebrafish: Regulation and Specification during the Determination of Neural Crest Cell-Fate.” *Pigment Cell & Melanoma Research* 16 (3): 172–82.

Zirlinger, Mariela, Liching Lo, Jill McMahon, Andrew P McMahon, and David J Anderson. 2002. “Transient Expression of the bHLH Factor Neurogenin-2 Marks a Subpopulation of Neural Crest Cells Biased for a Sensory but Not a Neuronal Fate.” *Proceedings of the National Academy of Sciences of the United States of America* 99 (12): 8084–89.

7

Appendices

List of Appendices

I. Plasmid maps

II. Primer sets

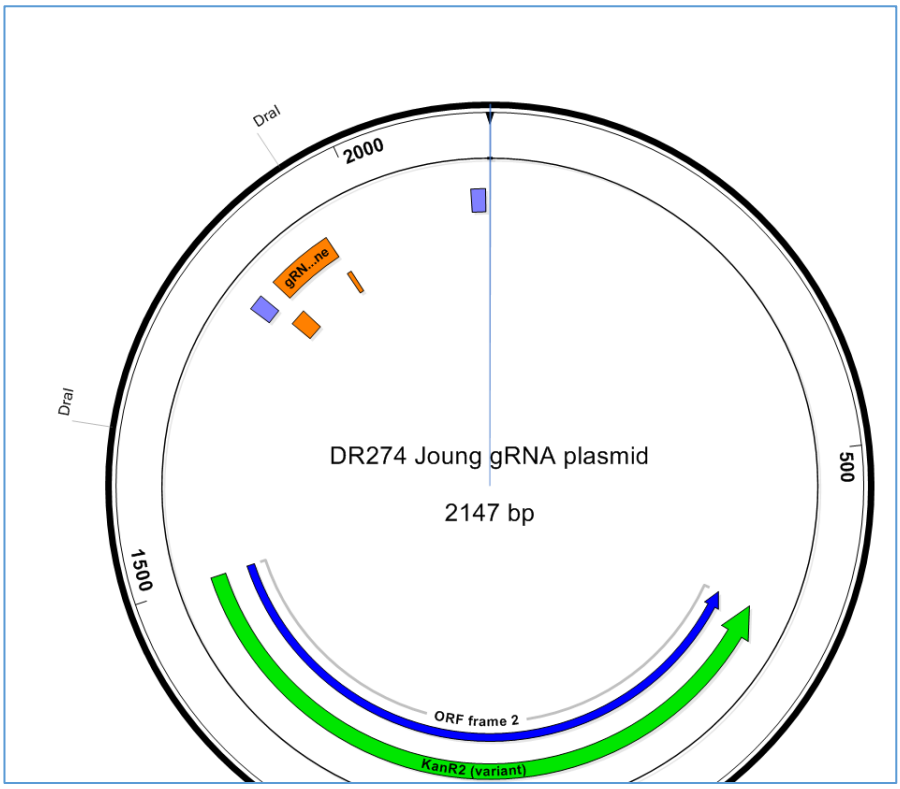
III. *foxd3*^{sa20726} genotyping

IV. MATLAB scripts for GRNs

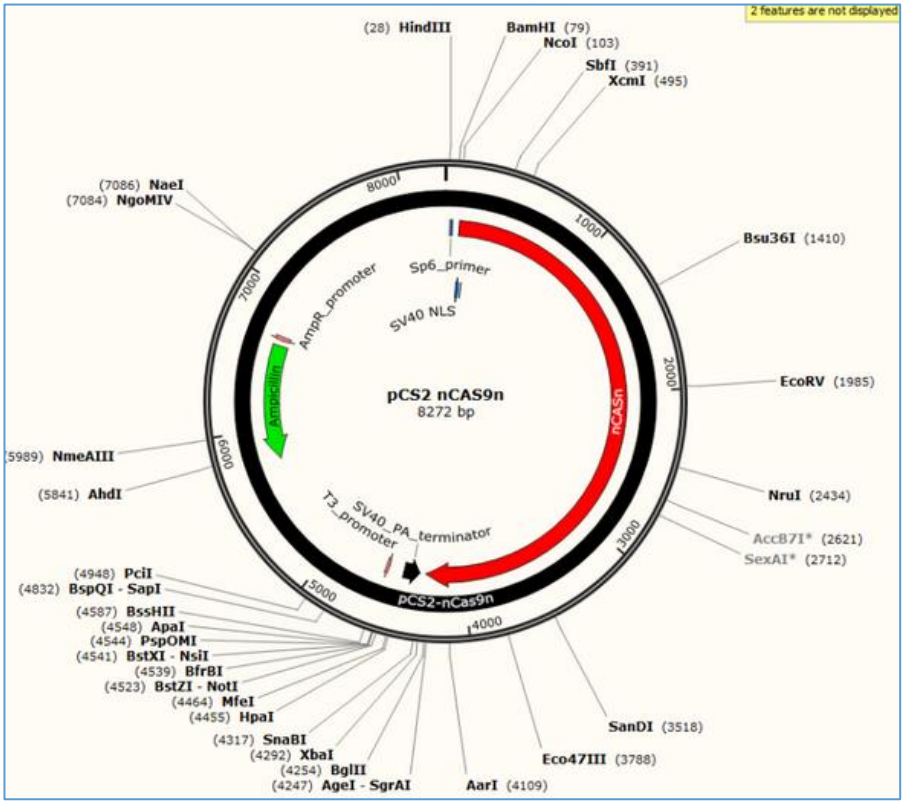
V. MATLAB scripts (Monte Carlo)

I. Plasmid maps

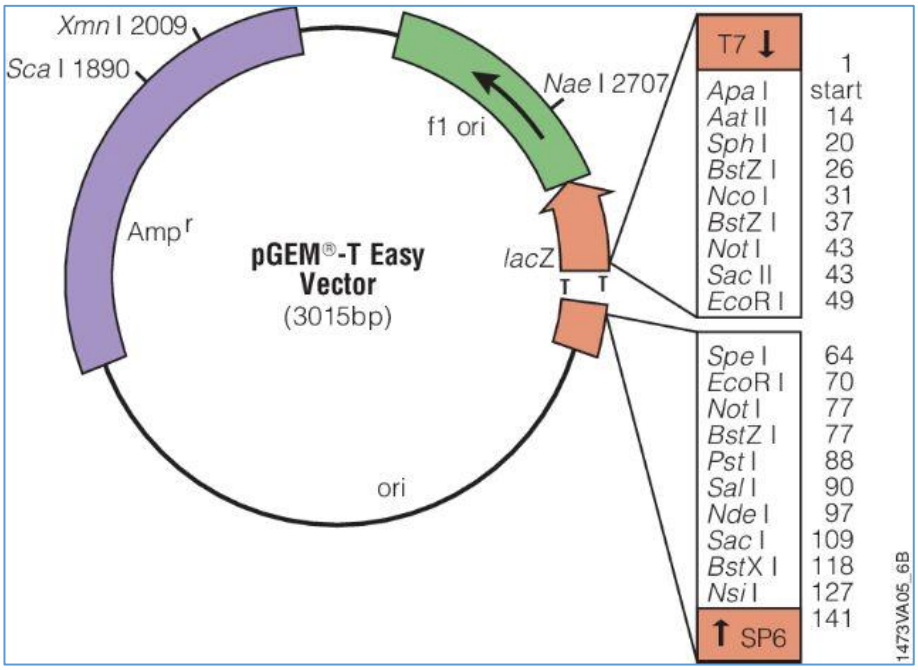
1. CRISPR gRNA sequence vector, obtained from the Sanger Institute. Purple boxes indicate the positions of the T7 promoter.



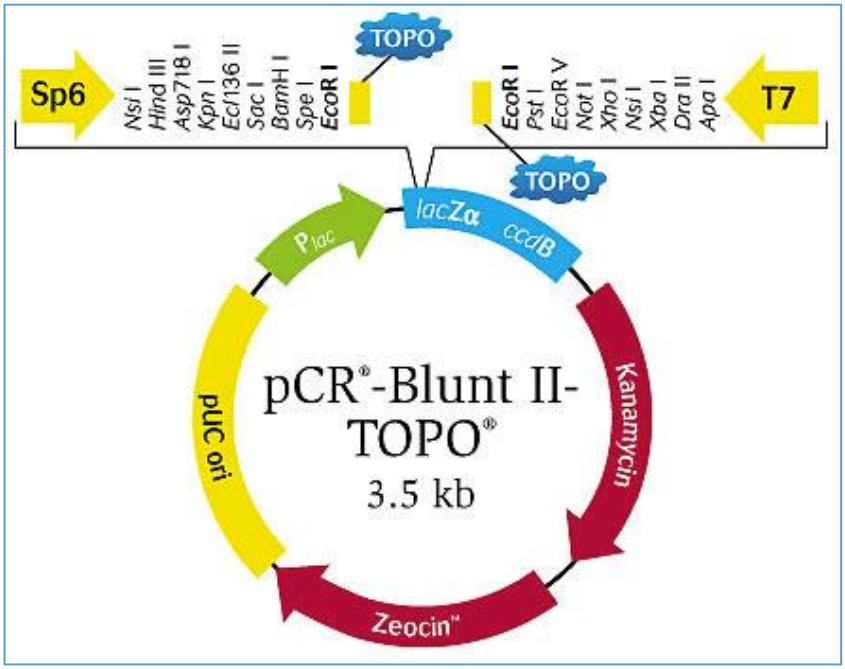
2. Cas9 sequence vector, obtained from the Sanger Institute.



3. pGEM-T Easy Vector (Promega)



4. pCR II-Blunt-TOPO Vector (Invitrogen)



II. Primer sets

Table II.1. Primer sets used for PCR-based genotyping.

Primer pair	Amplicon size	Allele	Notes
FW (1) 5'- TGA TGA CCC TGT CTG GAG G -3' **RV (1) 5'- CTT GGT CTC TCC GTC AGC G -3'**	227 bp	*foxd3*^{sa20726}	FW (1)/RV (1) primer set worked more efficiently than FW (1)/RV (2)
FW (1) 5'- TGA TGA CCC TGT CTG GAG G -3' **RV (2) 5'- CTG CTC TTG GGC TTG TTT TGC -3'**	280 bp	*foxd3*^{sa20726}	
FW 5'- CTA ACT CAA AGC AGT TTC GT -3' **RV 5'- GTA ACG TCA TGA GCA GAT AA -3'**	490 bp	*ltk*^{ty82}	(Lopes et al. 2008)

Table II.2. Primer sets used for HRMA-based genotyping.

Primer pair	Amplicon size	Allele	Conditions
FW 5'- ATA TCG ACG TGG TCG GGG -3' **RV 5'- TCA GCG TTG CTC TCG CAG GG -3'**	142 bp	*foxd3*^{sa20726}	2.5 mM MgCl2, 58.5 ºC
FW 5'- TAA GCC CAG TGA GGT CCT TCC -3' **RV 5'- TTT TTG TTC TGC GGG ATG CTC G -3'**	175 bp	*id2a* (CRISPR1)	1.5 mM, 60 ºC
FW 5'- CTG GAA CAA AGG GAC GAT CC -3' **RV 5'- TGG ATC CGT AGC TGG AGT CT -3'**	138 bp	*tfec* (CRISPR)	1.5 mM, 58.5 ºC Reverse primer designed by Dr J. Lister

Table II.3. Primer sets used for *in situ* hybridisation probe synthesis.

Primer pair	Amplicon size	Target gene	Notes
FW 5'- AGC CAA CAA TCA CGA CAG TG -3' **RV 5'- CCA ATA GAA ACG GGA GGT CA -3'**	946 bp	*tfec*	-
FW 5'- GTG ACA GTG CCT ACC GAA ACA -3' **RV 5'- ATG ATG GTG GAC GAG CGA AT -3'**	1.3 kb	*ltk*	Designed by Dr Susana Lopes (Kelsh group)

Table II.4. Primer sets used for PCR and sequencing when screening for CRISPR/Cas9 mutations.

Primer pair	Amplicon size	Allele	Notes
FW 5'- AGG CAA GGT AAT GTC CGA GA -3' **RV 5'- TGG ATC CGT AGC TGG AGT CT -3'**	310 bp	*tfec* (CRISPR)	Primers designed by Dr J. Lister.
Int.1 5'- GCG CGA TCT GCA GGT CCA GAA TA -3' **Int.2 5'- TCG CAA ACC AGC GAG TGA ACA AA -3'**	400 bp	*id2a* (CRISPR1)	Designed at the Sanger Institute
Ext.1 5'- AGG CTG GTT ATC GCC GAA TTG GA -3' **Ext.2 5'- AAA TGG GCG GAT CTT TGT CGC TC -3'**	517 bp	*id2a* (CRISPR1)	
Int.1 5'- GTG TCC TGC TGT CCT CTG TG -3' **Int.2 5'- ACA TCC AGA ACA CCC CTG AC -3'**	518 bp	*id2a* (CRISPR2)	
Ext.1 5'- GCA ACT TCA CAC ATG CAC ACT GC -3' **Ext.2 5'- CAG ATC GCG CTC GAC TCC AAT TC -3'**	708 bp	*id2a* (CRISPR2)	

Table II.5. Primer sets used for overexpression studies.

Primer pair	Amplicon size	Target gene	Notes
Sequence not known	Not known	*rlp13*	Commercially designed & synthesised (Primerdesign Ltd)
FW 5'-TGG AGA CTG CTG AAC GAG AC -3' **RV 5'- GCA GGC TCT TGT AAT GCG AT -3'**	199 bp	*sox10*	Designed by Dr Tatiana Subkhankulova
FW 5'- CTG GAC CAT GTG GCA AGT TT -3' **RV 5'- TGA GGT TGT GGT TGT CCT TCT -3'**	105 bp	*mitfa*	Commercially designed and synthesised (Invitrogen)
FW 5'-GGAGCTTGGATTGCATGGAG -3' **RV 5'-TTGATCAGCACCGTACACCT -3'**	89 bp	*tfec*	Designed by Dr Tatiana Subkhankulova
FW 5'- TCA GGA CTG TCA GTT GCA GTT T -3' **RV 5'- GTC CGG ATC TCT CCC AAA TC -3'**	119 bp	*ltk*	Designed by Dr Tatiana Subkhankulova
Sequence not known	Not known	*dct*	Commercially designed & synthesised (Primerdesign Ltd)
FW 5'-TGGATGCAGTTGGAATGAGT -3' **RV 5'-TTGACAGTCTCGTTGTCCTCA -3'**	127 bp	*pnp4a*	Designed by Dr Tatiana Subkhankulova

III. *foxd3*^{sa20726} genotyping

In the sequence shown below:

Bold: HRMA primer sites

Underlined: PCR genotyping primer sites (FW1/RV1 set)

[C/A]: SNP, C to A transition

//: *DdeI* cutting site

>foxd3 DNA sequence (chromosome:Zv9:6:31784202:31786000:-1)
AGTCCTTAAAGCCTGCTTTTCAAAGCATGTGTCATCTTGACTCTTCGATTCAAAGCCACG
CGTAAGTTTTGAGGAAATACAGTGGGAAAAGAGACCCAAGCAAGCCTTCGGACTGGAGAA
ATATTTTCTAAAGTTGAAGTTTTCCTGACAGGCCGGGACCGACCTGCATTTATAGCCACC
GGTCGTCAACTGATCTTACCTTGGGTTGCTCCAGCAGCAAAGTGCGCGGTGATGACCCTG
TCTGGAGGCACCAGTGCCAGCAACATGTCCGGTCAGACCGTGCTCACAGCTGACGATGTG
G**ATATCGACGTGGTCGGGG**AGGGTGACGAGGGGATGGAGCAGGACAGTGAC**//**TG[C/A]G
AAAGCCAGTGCATGCAGGACCGGGGAGATGAGGTGGAGGAGATCGAGGTGAAGGAGCG
CAGCGACAGT**CCCTGCGAGAGCAACGCTGA**CGGAGAGACCAAGGGGGATGCTCAGGAG
AGCTCCACCGGTCCCATGCAAAACAAGCCCAAGAGCAGCCTGGTAAAGCCGCCCTACTCG
TACATCGCCCTCATCACCATGGCCATCCTCCAGAGCCCGCAGAAGAAGTTGACGCTCAGT
GGAATCTGCGAGTTCATCAGCAACCGCTTCCCATACTACCGGGAGAAGTTTCCGGCCTGG
CAAAACTCCATTCGCCATAACTTGTCGCTCAACGACTGCTTCGTCAAGATCCCACGGGAAC
CGGGCAACCCGGGCAAAGGCAACTACTGGACCCTCGACCCCCAGTCGGAAGATATGTTC
GACAACGGTAGCTTTCTGAGGAGGAGAAAACGCTTCAAGAGGCATCAGCCGGACATTCTC
AGGGACCAGACCGCCCTCATGATGCAGAGTTTTGGGGCATACGGCATTGGGAATCCATAT
GGACGTCATTATGGAATTCACCCGGCTGCATACACGCACCCTGCCGCTCTGCAGTACCCG
TACATTCCCCCTGTGGGTCCGATGCTCCCTCCGGCGGTGCCTCTCTTACCCTCCGCCGAA
CTGAACAGA

foxd3^{sa20726} genotyping

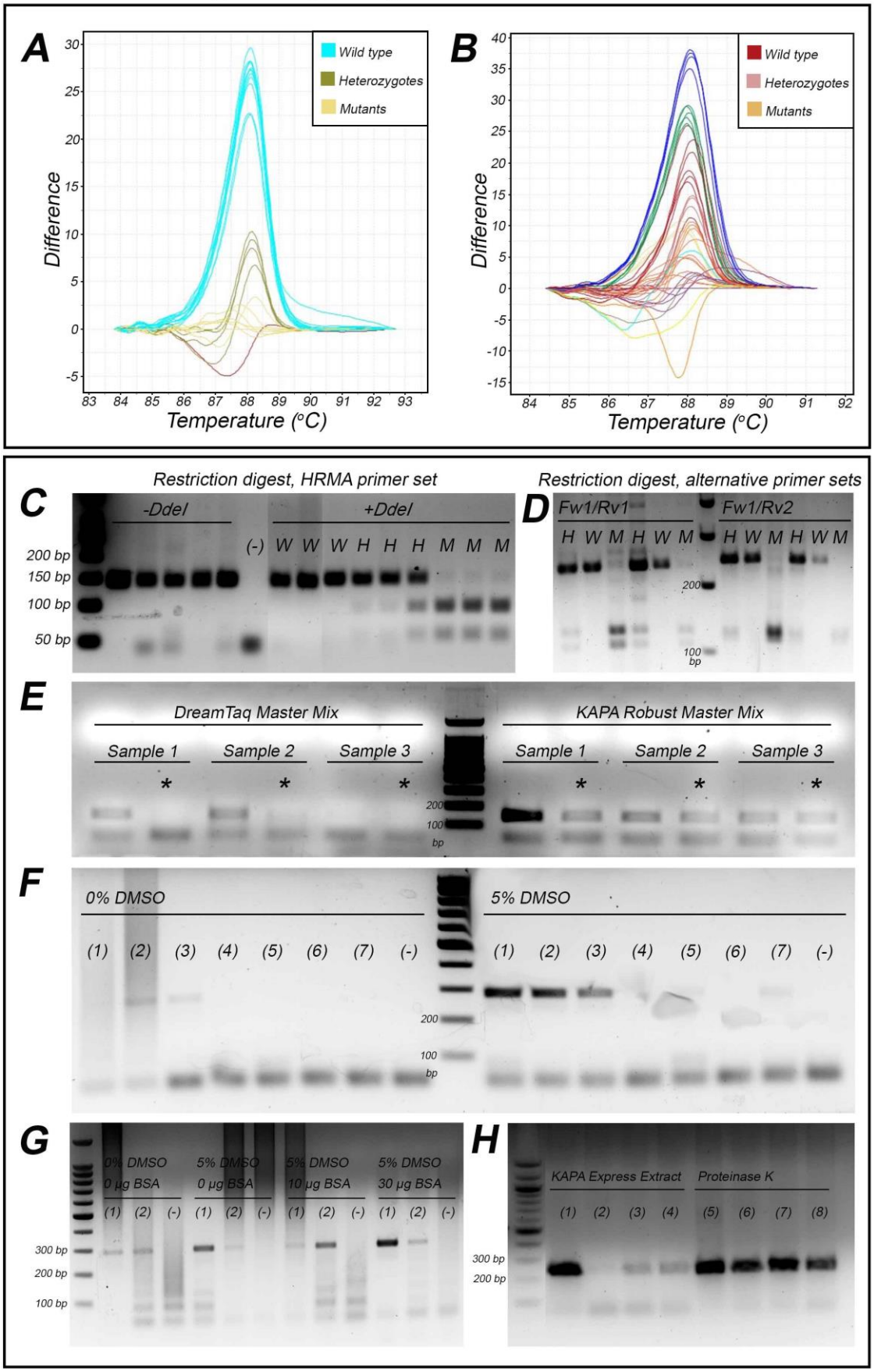


Figure III.1. *foxd3*^{sa20726} genotyping. HRMA-based genotyping using genomic DNA extracted from live embryos (A) and fixed embryos subjected to *in situ* hybridisation (B). In (B) several variants are predicted (different coloured curves) as a result of inefficient amplification. PCR and restriction digest-based genotyping using KAPA Robust DNA Polymerase and *DdeI* restriction enzyme with the HRMA primer set (C) and with newly designed primer sets producing larger amplicons in order to better differentiate distinct fragments (D). The amplification efficiency and yield were optimised so that fragments resulting from restriction digest can be discerned (E) by selecting the KAPA Robust DNA polymerase to perform the PCR reaction, (F) by adding 5% DMSO and (G) 30 µg BSA in the final reaction. A de-crosslinking step was attempted but did not optimise amplification (E, asterisks). (H) To extract genomic DNA, a proteinase-K extraction method was more efficient than the typically used KAPA Express Extract kit. In (H), (1) and (5) indicate samples from live embryos while (2-4) and (6-8) indicate samples from embryos subjected to ISH. W,WT; H, heterozygote; M, mutant; Fw, forward; Rv, reverse.

IV. MATLAB scripts for GRNs

Indicatively, the function files (IV.A; WT context) and (IV.B; loss of *sox10* function context), as well as the MATLAB script (IV.C) used to simulate the preliminary GRN (Model A) were as follows:

IV.A. Model A function file (WT context)

```
function dy = modelA_function_wt (t,G)

%% Variables
% G(1) = Sox10 protein concentration
% G(2) = Tfec protein concentration
% G(3) = Ltk protein concentration
% G(4) = Ltk ligand extracellular concentration
% G(5) = Concentration of ltk/ligand complexes on the membrane of a cell
% G(6) = Mitfa protein concentration
% G(7) = Pnp4a protein concentration

%% Constant parameters
K_d = [0.1 0.2 0.3]; % Equilibrium dissociation constant; Units nM
g = 0.2; % Maximal mRNA production rate; nM per hour
d = [0.1 0.25 0.5]; % Protein degradation rate; 1/h
k_on = 10^(-2)*3600; % on-rate for protein-protein interaction; 1/nM x 1/h
k_off = 3*k_on; % off-rate for protein-protein interaction; 1/h

% r(j): average number of occupied regulatory elements at equilibrium
r(1) = G(1)/(K_d(2)+G(1)); % sox10 upregulates tfec
r(2) = G(5)/(K_d(1)+G(5)); % ltk/lig upregulates tfec
r(4) = G(6)/(K_d(3)+G(6)); % mitfa upregulates pnp4a
r(5) = G(2)/(K_d(3)+G(2)); % tfec upregulates ltk
r(6) = G(1)/(K_d(2)+G(1)); % sox10 upregulates mitfa
r(7) = G(2)/(K_d(3)+G(2)); % tfec upregulates pnp4a

%% Defining the system of ODEs:
% dy(i) = d[G(i)]/dt
dy = zeros(7,1);
% sox10 input: sox10 is expressed at 18hpf in the NC domain
dy(1) = heaviside(t)*heaviside(22-t)*(g-d(3)*G(1))- heaviside(t-22)*heaviside(100-
t)*d(3)*0.35*G(1);
```

```
% tfec is regulated by sox10 and ltk/ligand complex
dy(2) = g*(r(1)*(1-r(2))+r(2))-d(3)*G(2);
% ltk is regulated by tfec
dy(3) = g*r(5)-d(2)*G(3);
% ltk ligand concentration is assumed to remain constant
dy(4) = 0;
% ltk/ligand complex production rate is the result of both ltk and ligand binding via an
AND gate
dy(5)= k_on*G(3)*G(4)-k_off*G(5);
% mitfa is regulated by sox10
dy(6)= g*r(6)-d(3)*G(6);
% pnp4a is regulated by tfec and mitfa
dy(7)= g*(r(7)*(1-r(4))+r(4))-d(1)*G(7);
```

IV.B. Model A function file (*sox10* mutant context)

```
function dy = modelA_ function_sox10(t,G)

%% Variables
% G(1) = Sox10 protein concentration
% G(2) = Tfec protein concentration
% G(3) = Ltk protein concentration
% G(4) = Ltk ligand extracellular concentration
% G(5) = Concentration of ltk/ligand complexes on the membrane of a cell
% G(6) = Mitfa protein concentration
% G(7) = Pnp4a protein concentration

%% Constant parameters
K_d = [0.1 0.2 0.3]; % Equilibrium dissociation constant; Units nM
g = 0.2; % Maximal mRNA production rate; nM per hour
d = [0.1 0.25 0.5]; % Protein degradation rate; 1/h
k_on = 10^(-2)*3600; % on-rate for protein-protein interaction; 1/nM x 1/h
k_off = 3*k_on; % off-rate for protein-protein interaction; 1/h

% r(j): average number of occupied regulatory elements at equilibrium
r(1) = 0; % sox10 cannot upregulate tfec
r(2) = G(5)/(K_d(1)+G(5)); % ltk/lig upregulates tfec
r(4) = G(6)/(K_d(3)+G(6)); % mitfa upregulates pnp4a
r(5) = G(2)/(K_d(3)+G(2)); % tfec upregulates ltk
r(6) = 0; % sox10 cannot upregulate mitfa
```

```
r(7) = G(2)/(K_d(3)+G(2)); % tfec upregulates pnp4a

%% Defining the system of ODEs:
% dy(i) = d[G(i)]/dt
dy = zeros(7,1);
dy(1) = heaviside(t)*heaviside(22-t)*(g-d(3)*G(1))- heaviside(t-22)*heaviside(100-
t)*d(3)*0.35*G(1); % sox10 input: sox10 is expressed at 18hpf in the NC domain
dy(2) = g*(r(1)*(1-r(2))+r(2))-d(3)*G(2); % tfec is regulated by sox10 and ltk/ligand
complex
dy(3) = g*r(5)-d(2)*G(3); % ltk is regulated by tfec
dy(4) = 0; % ltk ligand concentration is assumed to remain constant
dy(5)= k_on*G(3)*G(4)-k_off*G(5); % ltk/ligand complex production rate is the result of
both ltk and ligand binding via an AND gate
dy(6)= g*r(6)-d(3)*G(6); % mitfa is regulated by sox10
dy(7)= g*(r(7)*(1-r(4))+r(4))-d(1)*G(7); % pnp4a is regulated by tfec and mitfa
```

IV.C. Model A script

```
clear all;
close all;
clc;

%setting options-adjusting integration parameters for error control
%RelTol (relative tolerance) is a measure of the error relative to the size of each
solution component. Roughly, it controls the number of correct digits in all solution
components, except those smaller than thresholds AbsTol(i).
% AbsTol(i) is a threshold below which the value of the ith solution component is
unimportant. The absolute error tolerances determine the accuracy when the solution
approaches zero
options = odeset('RelTol',1e-4,'AbsTol',[1e-4 1e-4 1e-4 1e-4 1e-4 1e-4 1e-4]);
%options = odeset('RelTol',1e-6);

[T,Y] = ode45(@modelA_function_[wt/sox10],[18 100],[0.4 0.25 0 1 0 0 0],options);
%setting interval: 18hpf to 100hpf and initial values (gene expression in the posterior
trunk NCC domain at 18hpf)

%setting figure output properties; gcf = fig;
set(findall(gca,'type','axes'),'fontsize',18);
set(findall(gca,'type','text'),'fontSize',18);
linewidth = 2;
```

```
figure(1);
plot(T,Y(:,1),'-r',T,Y(:,2),'-g',T,Y(:,3),'-b',T,Y(:,6),'-m',T,Y(:,7),'-k','LineWidth',linewidth)
%,T,Y(:,5),'-m',T,Y(:,6),'-y'
axis([18 100 0 1.5])
legend('Sox10','Tfec','Ltk','Mitfa','Pnp4a')
xlabel('time (hpf)')
ylabel('Concentration (nM)')
saveas(gcf,'Figure1_all.m')
print('Figure1_all','-dtiff','-r700')
```

V. MATLAB scripts (Monte Carlo)

Indicatively, the function file (V.A) and script (V.B) used to run Monte Carlo for model D.1 were as follows:

V.A. Model A function file (WT context)

```
function dy = MonteCarlo_D1(t,G)

global a

% r(j): average number of occupied regulatory elements at equilibrium
r(1) = G(1)/(a(13)+G(1)); % sox10 upregulates tfec
r(2) = G(5)/(a(14)+G(5)); % ltk/lig upregulates tfec
r(4) = G(6)/(a(15)+G(6)); % mitfa upregulates pnp4a
r(5) = G(2)/(a(16)+G(2)); % tfec upregulates ltk
r(6) = G(1)/(a(17)+G(1)); % sox10 upregulates mitfa
r(7) = G(2)/(a(18)+G(2)); % tfec upregulates pnp4a
r(8) = G(1)/(a(19)+G(1)); % sox10 is required for pnp4a upregulation
r(9) = G(2)/(a(20)+G(2)); % tfec upregulates sox10
r(10) = G(2)/(a(21)+G(2)); % tfec represses mitfa

%% Defining the system of ODEs:
% dy(i) = d[G(i)]/dt
dy = zeros(7,1);
dy(1) = a(1)*r(9)-a(6)*G(1); % sox10 is regulated by tfec
dy(2) = a(2)*(r(1)*(1-r(2))+r(2))-a(7)*G(2); % tfec is regulated by sox10 and ltk/ligand complex
dy(3) = a(3)*r(5)-a(8)*G(3); % ltk is regulated by tfec
dy(4) = 0; % ltk ligand concentration is assumed to remain constant
dy(5)= a(11)*G(3)*G(4)-a(12)*G(5); % ltk/ligand complex production rate is the result of both ltk and ligand binding via an AND gate
dy(6)= a(4)*r(6)*(1-r(10))-a(9)*G(6); % mitfa is upregulated by sox10 and repressed by tfec
dy(7)= a(5)*r(8)*(r(7)-r(4)*r(7)+r(4))-a(10)*G(7); % pnp4a is regulated by tfec and mitfa
```

V.B. Monte Carlo script

```
clear all;
close all;
clc;
```

```
%% Declaring parameter set 'a' as global, in order for all functions to share a *single* copy of 'a'
global a
N = 10000; % Setting number of runs

% Setting parameter mimimum and maximum values
a_min = [10^(-1)*0.2 10^(-1)*0.2 10^(-1)*0.2 10^(-1)*0.2 10^(-1)*0.2 ... % minimum max gene expression rates set to 0.02 nM/hour
   10^(-1)*0.5 10^(-1)*0.5 10^(-1)*0.5 10^(-1)*0.5 10^(-1)*0.5 ... % minimum for protein degradation rates set to 0.05 1/h
   10^(-2)*3600 3*10^(-2)*3600 ... % kon and koff rates remain unchanged
   10^(-2)*0.2 10^(-2)*0.2 10^(-2)*0.2 10^(-2)*0.2 10^(-2)*0.2 10^(-2)*0.2 10^(-2)*0.2 10^(-2)*0.2 10^(-2)*0.2... % minimum  Kds set to 0.002 nM
   0.01 0.01 0 0 0 0 0]; % minimum  initial values (nM)
a_max = [10^2*0.2 10^2*0.2 10^2*0.2 10^2*0.2 10^2*0.2 ... % maximum max gene expression rates set to 20 nM/hour
   10*5 10*5 10*5 10*5 10*5 ... % maximum protein degradation rates set to 50 1/h
   10^(-2)*3600 3*10^(-2)*3600 ... % kon and koff rates remain unchanged
   10^1*0.2 10^1*0.2 10^1*0.2 10^1*0.2 10^1*0.2 10^1*0.2 10^1*0.2 10^1*0.2 10^1*0.2... % maximum Kds set to 2 nM
   10 10 0 10 0 0 0]; % maximum initial values (nM)

S = zeros(N,29); % Defining matrix S in which scores and respective th values for each draw are categorised.

%% Setting up a 'for loop'; used to repeat specified number of times
for n=1:N

   % Randomising and scaling parameters
   th = rand(1,28);

   a = a_min+th.*(a_max-a_min); % Parameters are scaled using their assigned min and max values, as well as the new parameter vector th.

   % setting options-adjusting integration parameters for error control
   options = odeset('RelTol',1e-4,'AbsTol',[1e-4 1e-4 1e-4 1e-4 1e-4 1e-4 1e-4]);
   %options = odeset('RelTol',1e-6);
```

```
        % Calling function with differential equations
        [T,Y] = ode45(@MonteCarlo_D1,[18 100],[a(22) a(23) a(24) a(25) a(26) a(27)
a(28)],options);

        % Scoring based on maximum difference of maximum Mitfa concentration and
minimum positive value reached after the maximum
    [M,I] = max(Y(:,6)); % Finding maximum value M at time I (index)
    tail = 0; % Setting tail to 0, to clear values from previous runs
    tail = Y(I:end,6); % Defining 'tail' vector as the mitfa concentration values from I to
the end
    [rowMinTail, colMinTail] = find(tail==min(tail(tail>0)));   % Finding the minimum
positive value
    score = M/tail(rowMinTail(1), colMinTail(1)); % Scoring based on first value returned
    S(n,1) = score; % Categorising the score of each draw n in the first column of vector
S
    S(n,2:end) = th; % Categorising the respective th values of each draw n in the
second to last columns of vector S
end

%% Sorting the rows in descending order using the values in column 1. Saving sorted
S table as .txt
S_sorted = sortrows(S,-1); %Minus sign indicates descending order.
save('bestScores.txt','S_sorted','-ascii');

%% Running simulation and plotting curve for 'best th' according to score
best_th = S_sorted(1,2:end);

a = a_min+best_th.*(a_max-a_min);

options = odeset('RelTol',1e-4,'AbsTol',[1e-4 1e-4 1e-4 1e-4 1e-4 1e-4 1e-4]);
[T,Y] = ode45(@MonteCarlo_D1,[18 100],[a(22) a(23) a(24) a(25) a(26) a(27)
a(28)],options);

%setting figure output properties; gcf = fig;
set(findall(gca,'type','axes'),'fontsize',18);
set(findall(gca,'type','text'),'fontSize',18);
linewidth = 2;
```

```
figure(1);
plot(T,Y(:,1),'-r',T,Y(:,2),'-g',T,Y(:,3),'-b',T,Y(:,6),'-m',T,Y(:,7),'-k','LineWidth',linewidth)
%,T,Y(:,5),'-m',T,Y(:,6),'-y'
axis([0 100 0 inf])
legend('Sox10','Tfec','Ltk','Mitfa','Pnp4a')
xlabel('time (hpf)')
ylabel('Concentration (nM)')
%saveas(gcf,'Figure1_all.m')
print('Figure1_MC_D','-dtiff','-r700')
```